# Molecular studies of *Wolbachia* and sex-determination genes in Australian *Bactrocera* species – complementary approaches to improved fruit fly control

Jennifer Louise Morrow

BSc (Hons)

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Hawkesbury Institute for the Environment

University of Western Sydney, Hawkesbury Campus

June 2014

# Acknowledgements

First to my principal supervisor Markus Riegler, who has guided and encouraged me in my work over the last four years, gone over and above in time and effort, received and responded to emails at all hours, and provided great opportunities to learn and research in a supportive environment.

I am indebted to my supervisors Marianne Frommer and Deborah Shearman, whom I have known for many years. They gave me a wonderful gift by offering a way back into science after a seven year break, and showed their belief in me by their encouragement to embark on a PhD. Thank you for seeing it through with me.

I acknowledge with gratitude the UWS Hawkesbury Foundation, for granting me the F.G Swain Award, which enabled me to undertake transcriptome sequencing; the Australian Academy of Science, whose generosity enabled me to attend an EU COST 850 bioinformatics workshop in Sweden and a visit to the laboratory of Ass. Prof. Wolfgang Miller in Vienna; and the Genetics Society of Australasia for the Smith-White Travel Award which allowed me to attend the International Congress of Entomology 2012 in Korea. I am also grateful for receiving the Australian Postgraduate Award and top-up scholarships from the University of Western Sydney and the Hawkesbury Institute for the Environment.

Many thanks go to those who have shared their knowledge and skills over the course of my candidature. I am particularly grateful to Stuart Gilchrist, John Sved, Paul Rymer, Remko Duursma, Jeff Powell and Barbara Drigo for advice on statistics and bioinformatics, Goran Lopaticki for excellent work looking after the fly stocks, Caroline Janitz for the Next-Generation sequencing and Wolfgang Miller and Daniela Schneider for teaching me new techniques and welcoming me to Vienna.

Sincere thanks to the students who were here when I first joined UWS, especially Megan Halcroft, Jocelyn King and Jenny Shanks, who helped me in getting settled, befriended me, showed me the way things worked and co-ordinated our lunch-breaks. To Duong Nguyen, Aidan Hall, Andrew Gherlenda, Tim Sutton and Peter Kern, it’s been a pleasure sharing the molecular and entomology labs with you, thanks for everything.

On a personal note, there are so many people who have encouraged me over the last four years. First to my mum and dad, June and Vince Bayliss who have been the best parents one could ever hope for. I am so grateful for their love, encouragement and practical support, including hours of school holiday child-care and after school pick-ups. To Michelle Page, for being a supportive friend and always being on hand to help in whatever way was needed. To Sandy, Mary, Deb, Anne and Liz for your encouragement and prayers, and to Lionel for being my unofficial mentor.

Finally, my love, thanks and big hugs go to John and our three amazing children Christopher, Andrew and Amy. I couldn’t have done this without you all being so helpful, easy-going and patient, and for always making me smile.

# Statement of authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

..............................................

Jennifer Morrow

June 2014

# Table of contents

**List of tables** iv

**List of figures** vii

**Abbreviations** x

**Abstract** xiii

**Chapter 1 General Introduction** 1

**1.1 Biology of fruit flies** 2

**1.2 Fruit fly pest control strategies** 6

**1.3 Symbiotic bacteria in insects** 10

**1.4 *Drosophila melanogaster* somatic sex-determination pathway** 16

**1.5 The somatic sex-determination pathway in Tephritidae** 24

**1.6 Research scope and aims** 29

**1.7 Thesis structure** 32

**Chapter 2 Latitudinal cline in the incidence of *Wolbachia* endosymbionts in Australian tephritid fruit fly communities** 33

**2.1 Abstract** 34

**2.2 Introduction** 35

**2.3 Materials and methods** 37

**2.4 Results** 41

**2.5 Discussion** 50

**2.6 Conclusions** 54

**Chapter 3 Tropical tephritid fruit fly community with high incidence of shared *Wolbachia* strains as platform for horizontal transmission of endosymbionts** 55

**3.1 Summary** **56**

**3.2 Introduction** **57**

**3.3 Results** **59**

**3.4 Discussion** **67**

**3.5 Conclusions** **74**

**3.6 Experimental procedures** **75**

**Chapter 4 The microbiome of field-caught and laboratory-adapted Australian tephritid fruit fly species with different host plant use and specialisation** **81**

**4.1 Abstract** **82**

**4.2 Introduction** **83**

**4.3 Materials and methods** **86**

**4.4 Results** **89**

**4.5 Discussion** **100**

**Chapter 5 Expression patterns of sex-determination genes in single male and female embryos of two *Bactrocera* fruit fly species during early development** **109**

**5.1 Abstract** **110**

**5.2 Introduction** **111**

**5.3 Results** **113**

**5.4 Discussion** **125**

**5.5 Conclusions** **129**

**5.6 Experimental procedures** **130**

**Chapter 6 Transcriptome sequencing of male and female *Bactrocera jarvisi* embryos** **137**

**6.1 Abstract** **138**

**6.2 Background** **139**

**6.3 Methods** **142**

**6.4 Results** **145**

**6.5 Discussion** **159**

**6.6 Conclusion** **162**

**Chapter 7 General Discussion** **164**

**7.1 Introduction** **165**

**7.2 Key findings and future applications** **167**

**7.3 General conclusions** **184**

**References** **185**

**Appendices** **222**

**Appendix A** **223**

**Appendix B** **249**

**Appendix C** **260**

**Appendix D** **265**

**Appendix E** **271**

# List of tables

**Table 1.1** References for the primary sex-determination genes in tephritids, *D. melanogaster* and the house fly, *M. domestica.* 24

**Table 2.1** *Wolbachia* prevalence in six Australian fruit fly species across climate zones. 42

**Table 2.2** Multivariate and univariate analyses of *Wolbachia* abundance at different latitudes. 45

**Table 2.3** *wsp* allele numbers and GenBank accession numbers from species, confirmed in listed individuals by sequencing and PCR-RFLP. 46

**Table 3.1** MLST and *wsp* alleles from nine tephritid species and one parasitoid species *Fopius arisanus.* 61

**Table 3.2** Parasitoid wasps used in this study. 63

**Table 3.3** Larvae collected from wild tobacco plants in Atherton, Queensland identified as *B. cacuminata.* 64

**Table 4.1** Alpha diversity metrics of 454 pyrosequencing of twelve samples, calculated at the 97% identity level. 90

**Table 4.2** Percentage abundance of family and genus level taxa from each sample, based on RDP Classification. 95

**Table 5.1** *tra* female-specific amino acid percentage identity between species, expressed as a fraction of total non-gap sites only. 115

**Table 5.2** *tra-2* amino acid percentage identity (below diagonal) and fraction of differences (above diagonal) between species. 117

**Table 5.3** *Sxl* amino acid percentage identity (below diagonal) and fraction of differences (above diagonal) between species. 117

**Table 6.1** Results of trimming and filtering poor quality reads in CLC Genomics Workbench. 147

**Table 6.2** CLC *de novo* assembly of transcripts from all eight libraries. 147

**Table 6.3** Trinity *de novo* assembly of sequences from all eight *B .jarvisi* libraries. 148

**Table 6.4** Expression profiles of the 19 sex-determination genes identified in the *B. jarvisi* CLC assembly. 151

**Table 6.5** *Bactrocera jarvisi* cellularisation gene homologues. 154

**Table 6.6** Summary of differential expression experiments performed in CLC Genomics Workbench. 156

**Table 6.7** Differential expression results showing significantly up-regulated transcripts in males over females in 3-5h samples (Experiment B) and 2-3h samples (Experiment D). 158

**Table A.1** Collection data for field-caught fruit flies from this study. 223

**Table A.2** Laboratory strains screened for *Wolbachia* using *wsp* and 16S rRNA loci. 243

**Table A.3** PCR conditions and primer sequences used in Chapter 2. 245

**Table A.4** Comparison of *Wolbachia* infected individuals sampled from equatorial and tropical regions for **(A)** *B. bryoniae*, **(B)** *B. frauenfeldi*, **(C)** *B. neohumeralis,* **(D)** *B. strigifinis* and **(E)** *B. tryoni* over five collection years. 247

**Table A.5** Distance matrix of *COI* sequences of *Bactrocera tryoni* and *Bactrocera neohumeralis* individuals. 248

**Table B.1** Collection data for 104 field-caught, male fruit flies from this study. 249

**Table B.2** *Wolbachia*, mitochondrial and qPCR primers. 253

**Table B.3** Number of *Wolbachia* isolates in the MLST database that share MLST alleles with ST-17 and ST-370 isolated from *B. frauenfeldi, B. cacuminata* and *F. arisanus*. 254

**Table B.4** *Wolbachia* strain density values. 255

**Table B.5** ANOVA of the linkage of genes within each *Wolbachia* chromosome. 257

**Table B.6** PCR conditions for *Wolbachia* detection, MLST characterisation, quantification and *COI* barcoding. 258

**Table B.7** Primer efficiency (E) values calculated from standard curves for host nuclear gene *scarlet*, and *Wolbachia* allele-specific primers. 259

**Table C.1** Field and laboratory fruit fly samples. 260

**Table C.2** Concentration and quality of pooled sample DNA for 454 pyrosequencing. 262

**Table C.3** Quality statistics for sequencing reads for 12 fruit fly samples. 263

**Table C.4** Jaccard distance matrix between samples, implemented in RDP II. 264

**Table C.5** UniFrac distance matrix between samples, unweighted distances below the diagonal, weighted distances above the diagonal. 264

**Table D.1** Primer sequences used in Chapter 5 265

**Table D.2** PCR conditions and primer sequences used in Chapter 5. 268

**Table D.3** qRT-PCR primers used in this study. 269

**Table D.4** qRT-PCR efficiency values for each primer pair. 270

**Table E.1** PCR conditions and primer sequences used in Chapter 6. 271

**Table E.2** Quality and concentration of total RNA samples for transcriptome sequencing. 272

**Table E.3** Sex-determination and cellularisation genes. 273

**Table E.4** A list of the 66 contigs up-regulated in older male embryos, after exclusion of contigs also up-regulated in older females. 276

**Table E.5** A list of the *Bactrocera jarvisi* contigs up-regulated in older embryos with BLAST matches to *C. capitata* and *Drosophila* genes. 279

# List of figures

**Figure 1.1** Morphology and distribution of the major horticultural pest species, *Bactrocera tryoni*. 3

**Figure 1.2** Morphology and distribution of *Bactrocera neohumeralis*. **5**

**Figure 1.3** Morphology and distribution of *Bactrocera jarvisi*. **6**

**Figure 1.4** Diagram showing the four main reproductive phenotypes caused by *Wolbachia* in different arthropods. 12

**Figure 1.5** Expression of *Drosophila melanogaster Sxl* gene from the early (*Sxl*Pe) and late (*Sxl*Pm) promoters. 19

**Figure 1.6** Simplified sex-determination pathway in *Drosophila melanogaster*, where alternative splicing of mRNA leads to different protein products in males and females. 21

**Figure 2.1** Incidence of *Wolbachia* infection in tephritid species, drawn on the Köppen map of Australia, with climate zones and latitudes as approximate climate divisions along the east coast of Australia. 44

**Figure 2.2** Bayesian inference tree of *wsp* sequences from *Wolbachia* within tephritid flies and other invertebrates. 47

**Figure 2.3** Bayesian inference tree of *COI* sequences from representatives of the five major tephritid fruit fly genera, *Anastrepha, Bactrocera*, *Ceratitis*, *Dacus* and *Rhagoletis* as well as *Dirioxa.* 49

**Figure 3.1** Comparison of *wsp* alleles' overall titres, and relative difference of strain titre within and across species based on *wsp* alleles using qPCR. 66

**Figure 3.2** Relative titre of *wsp*, *fpbA* and *coxA* alleles shared between singly and doubly-infected individuals of six tephritid species. 68

**Figure 3.3** Bayesian inference phylogenetic tree derived from concatenated MLST DNA sequences from this study and selected sequence types from the *Wolbachia* MLST database. 69

**Figure 4.1** Rarefaction curves of OTUs clustered at 97% similarity for each tephritid fruit fly sample. 90

**Figure 4.2** PCoA of Jaccard distances based on OTUs calculated at 97% similarity. 91

**Figure 4.3** PCoA of unweighted UniFrac distances based on OTUs calculated at 97% similarity. 92

**Figure 4.4** PCoA of weighted UniFrac distances based on OTUs calculated at 97% similarity. 93

**Figure 4.5** Relationships between hosts and microbial composition. The heatmap shows the relative abundance of the 15 most common OTUs, determined at 95% similarity. 94

**Figure 4.6** Relative abundance of bacterial families represented by 16S rDNA reads. 99

**Figure 5.1** Genetic structure of the *transformer* gene of *B. tryoni* **(A)** and *B. jarvisi* **(B)**. 114

**Figure 5.2** The genetic structure of the *tra-2* gene in *B. tryoni* and *B. jarvisi* 116

**Figure 5.3** Partial genetic and transcript structure of the *Sxl* gene in *B. tryoni* and *B. jarvisi.* 118

**Figure 5.4** qRT-PCR relative transcript abundance for BJ male and female embryos aged from 1-9h. 121

**Figure 5.5** qRT-PCR relative transcript abundance for BTJ[Y] male and female embryos aged from 1-9h. 122

**Figure 5.6** *B. jarvisi*, *B. tryoni* and BTJ[Y] embryos at different stages of early development 124

**Figure 6.1** Species distribution of BLAST2GO top matches against the Trinity CDS assembly 148

**Figure 6.2** Classification of the gene ontology (GO) terms for the *B. jarvisi* transcriptome. 149

**Figure 6.3** Principal co-ordinates analysis of the eight transcriptome samples mapped to the CLC assembly. 150

**Figure 7.1** Diagram depicting transgenic and non-transgenic methods for improving the current pest management repertoire, and the associated advantages and disadvantages. 179

**Figure A.1** Relationship between prevalence of *Wolbachia* infection in tephritid species and the midpoint of the latitudinal distribution 246

.

# Abbreviations

| | |
|---|---|
| aa | amino acids |
| ANOVA | analysis of variance |
| bHLH | basic helix-loop-helix |
| *bnk* / BNK | *bottleneck* gene / protein |
| bZIP | basic leucine zipper domain |
| CI | cytoplasmic incompatibility |
| *COI* | *cytochrome oxidase I* |
| *coxA* | *cytochrome oxidase subunit A* gene |
| Cq | quantification cycle |
| *da* / DA | *daughterless* gene / protein |
| DIG | digoxygenin |
| *dpn* / DPN | *deadpan* gene / protein |
| *dsx* / DSX | *doublesex* gene / protein |
| *emc* / EMC | *extra-machrochaetae* gene / protein |
| ESE | exonic splicing enhancer |
| *fbpA* | *fructose-bisphosphate aldolase* gene |
| *fru* / FRU | *fruitless* gene / protein |
| *ftsZ* | *cell division protein* gene |
| *gatB* | *aspartyl/glutamyl-tRNA amidotransferase subunit B* gene |
| GMO | genetically modified organism |
| GTR | general time reversible model |

| | |
|---|---|
| *gro* / GRO | *groucho* gene / protein |
| *hcpA* | *conserved hypothetical protein gene of Wolbachia* |
| *her* / HER | *hermaphrodite* gene / protein |
| IIT | incompatible insect technique |
| *ix* / IX | *intersex* gene / protein |
| *M* | *Dominant Male Determiner* |
| MLST | multi-locus sequence typing |
| MVLA | multi-locus variable number tandem repeat analysis |
| MZT | maternal-to-zygotic transition |
| OD1 / OD2 | oligomerisation domain 1 / 2 |
| ORF | open reading frame |
| *os* / OS | *outstretched* gene / protein |
| PCR-RFLP | polymerase chain reaction-restriction fragment length polymorphism |
| PLG | phase lock gel (tubes) |
| PRE | purine-rich elements |
| qPCR | quantitative PCR |
| qRT-PCR | quantitative reverse transcription PCR |
| RBD | RNA binding domain |
| *Rbp-1 /* RBP-1 | *RNA-binding protein 1* gene / protein |
| RIDL | release of insects carrying a dominant lethal |
| RRM | RNA-recognition motif |
| RS | arginine / serine |

| | |
|---|---|
| *run* / RUN | *runt* gene / protein |
| *sc* / SC | *scute* gene / protein |
| *sisA* / SISA | *sisterless A* gene / protein |
| SIT | sterile insect technique |
| *slam* / SLAM | *slow as molasses* gene / protein |
| SR | splicing regulator |
| *sryα* / SRYα | *serendipityα* gene / protein |
| ST | sequence type |
| *Sxl* / SXL | *Sex-lethal* gene / protein |
| *Sxl*Pe | *Sex-lethal* establishment (early) promoter |
| *Sxl*Pm | *Sex-lethal* maintenance (late) promoter |
| TN93 | Tamura - Nei |
| *tra* / TRA | *transformer* gene / protein |
| *tra-2* / TRA-2 | *transformer-2* gene / protein |
| U2AF | U2 auxiliary factor |
| *wsp* | *Wolbachia surface protein* |
| XSE | X-linked signal element |

# Abstract

*Bactrocera tryoni* (Diptera, Tephritidae), the Queensland fruit fly (Qfly), is the most significant horticultural pest species in Australia. The Sterile Insect Technique (SIT), which relies on the release of vast numbers of irradiated sterile insects to suppress field populations, has been used in this species and, globally, is a successful pest management strategy for many insect species. For Mediterranean fruit fly, *Ceratitis capitata*, the incorporation of a genetic sexing system (GSS), that allows male-only releases, can increase effectiveness of SIT by 3-5 times. However, there is currently no GSS for *B. tryoni*. Furthermore, field performance of sterile flies, mass-reared and irradiated, can be generally lower than of wild individuals.

Novel strategies such as Incompatible Insect Technique (IIT) exploit the reproductive manipulations of the maternally-inherited common endosymbiotic bacterium of insects, *Wolbachia*, to induce unidirectionally incompatible matings between uninfected field-females and released males carrying *Wolbachia*. While this method circumvents irradiation, a male-only release cohort is essential because females carrying *Wolbachia* are fully fertile; accidental release of infected females and thus inheritance of released *Wolbachia* in field populations would soon lead to a breakdown of the mating incompatibility essential for pest suppression.

Introductory Chapter 1 introduces the conceptual framework and background information for the thesis to then deliver the specific research questions, aims and thesis structure. Chapter 2 details the incidence and prevalence of *Wolbachia* in Australian fruit flies, to provide the initial, essential assessment of the feasibility of IIT in controlling Australian pest fruit flies. A field survey of 24 tephritid fruit fly species, collected along a continuing gradient through four climate zones of eastern Australia, with PCR screening of *Wolbachia* surface protein gene (*wsp*) and 16S rDNA, detected *Wolbachia* in eight species. Unexpectedly, the incidence of *Wolbachia* in fruit fly communities was restricted to northern Australia, including infections in five fruit fly species that only occur in the tropics. *Wolbachia* prevalence within three widely-distributed tephritid species, including the two economically important pests, *B. tryoni* and *Bactrocera neohumeralis*, was also

highest in the tropics. While this is the first potential demonstration of latitude as a contributing factor in the occurrence of *Wolbachia* in insect communities, the overall low infection frequency, particularly in *B. tryoni*, advocates for further research and development of *Wolbachia*-based control strategies.

To further characterise the *Wolbachia* strains found in tephritids from the more tropical regions of Australia, Multilocus Sequence Typing (MLST) and a novel quantitative PCR method for allele assignment in multiply-infected individual flies were applied, as described in Chapter 3. Based on five MLST loci and the *Wolbachia* surface protein gene (*wsp*), individuals of five *Bactrocera* and one *Dacus* species were found to harbour the same two strains as double infections, and *B. tryoni* harboured one of those strains. The sharing of *Wolbachia* was also evident across different trophic levels, as a fruit fly parasitoid species, *Fopius arisanus* (Hymenoptera: Braconidae), shared identical alleles with two *Wolbachia* strains detected in *Bactrocera frauenfeldi*. As *Wolbachia* are primarily maternally inherited, with only occasional documented horizontal transmission, this tropical insect community demonstrates an unprecedented, high incidence of four shared *Wolbachia* strains in eight host species from two trophic levels. Tropical insect communities, as studied here, may act as horizontal transmission platforms that contribute to the ubiquity of the otherwise maternally-inherited *Wolbachia.*

The fitness of insects produced by mass-rearing facilities for both SIT and IIT must ensure competitiveness with wild males upon release. The nutritional and physiological benefits derived from bacterial associations other than *Wolbachia* may be an important ancillary to optimise the fitness of flies reared for mass-release, both before and after sterilisation. Therefore, in Chapter 4, the bacterial communities associated with whole fruit flies from laboratory and natural fly populations were explored by 454 pyrosequencing of 16S rDNA amplicons, to reveal dominant and rare bacterial taxa. Six tephritid fruit fly species from three genera were investigated, including species that are polyphagous pests (*B. tryoni, B. neohumeralis, Bactrocera jarvisi*, *C. capitata*) and monophagous specialists (*Bactrocera cacuminata*). These were compared with the microbiome of a non-pestiferous but polyphagous tephritid species that is restricted to damaged or rotting fruit (*Dirioxa pornia*). Comparative analyses based on presence / absence and phylogenetic distance metrics discerned higher diversity in polyphagous over monophagous species and in field-collected

over laboratory-reared individuals; demonstrating that ecological niche and laboratory environment appeared to be important drivers of bacterial community composition and relative abundance. Overall, bacterial composition in Australian tephritids was dominated by Enterobacteriaceae and Acetobacteraceae (Proteobacteria), and Streptococcaceae and Enterococcaceae (Firmicutes).

A significant drawback in the application of SIT and development of IIT for *B. tryoni* in Australia has been the lack of a GSS. Chapter 5 describes the identification of gene homologues involved in sex determination that may be applied to a transgenic or RNAi-based sexing system to generate the necessary molecular targets for the production of a male-only strain. Analysis of expression patterns of these genes in early embryonic stages was performed to determine the important developmental time of expression of the as-yet uncharacterised male determining locus (*M*) after fertilisation. As detailed in Chapter 6, transcriptome sequencing at this specified time was carried out to identify candidate transcripts for *M* and other early expressed transcripts, as possible sources of regulatory sequences for transgenic sexing systems.

Analyses were carried out on two species, *B. tryoni* and *B. jarvisi*. Both explicitly demonstrated the conservation of gene and transcript structure for genes involved in sex-determination: *transformer* (*tra)*, *transformer-2* (*tra-2)* and *Sex-lethal* (*Sxl*). A single, sexed embryo approach employing qRT-PCR was used to examine the time course of expression of these genes in the sex-determination pathway in individual male and female *Bactrocera* during early embryonic development. Embryo sexing of a *B. jarvisi* line plus a *B. tryoni* line with an introgressed *B. jarvisi* Y-chromosome was possible due to molecular markers located on the *B. jarvisi* Y-chromosome. Analysis by qRT-PCR showed that, in *B. jarvisi*, the uncharacterised male determining locus, *M,* must have actively influenced expression of sex-specific *tra* transcripts within 6 hours of egg laying. Therefore, *B. jarvisi* embryos, collected 2-3h and 3-5h after egg laying, were individually sexed using the Y-chromosome genetic marker, and the poly(A)$^+$ transcriptome sequenced and assembled *de novo.* Fifteen sex-determination gene homologues and two cellularisation gene homologues of *Drosophila melanogaster* (Diptera: Drosophilidae) were identified in *B. jarvisi*; no candidates for *M* were identified in the poly(A)$^+$ mRNA fraction, and may instead be found through sequencing the non-protein-coding RNAs.

This thesis details the results of three fundamental and independent, yet complementary, experiments designed to investigate prospects for IIT and improvement of SIT, namely:

- an assessment of the incidence, prevalence and type of *Wolbachia* strains in Australian tephritid fruit fly species across their geographic range;
- a survey of the general microbiome of tephritid fruit fly species from different ecological niches and both natural and laboratory sources;
- an examination of genes expressed in early embryos of two Australian *Bactrocera* pest species, to generate molecular tools for GSS.

Moreover, this work represents a valuable contribution to the ecological and epidemiological understanding of *Wolbachia* infection life cycles in insect communities, through discovery of a new model system for examining the distribution of *Wolbachia* along a climatic gradient, and the horizontal transmission of *Wolbachia* among a community of tephritid fruit flies. This work also presents important molecular studies of sex-specific gene expression over early embryonic development, with the aim of clarifying the function of *M* in tephritid fruit flies as models for the many insect species with similar sex-determination systems.

# Chapter 1

# General Introduction

## 1.1 Biology of fruit flies

**"The principal components of the life systems of tephritids are moisture, temperature, light, food, natural enemies, and symbiotes."**

**M. A. Bateman, 1972**

Tephritid fruit flies, also called true fruit flies, are a widespread group found throughout tropical, subtropical and temperate regions of the world (Christenson and Foote 1960). There are approximately 5,000 species defined in the family Tephritidae, and the females feature a long extendable ovipositor, used to lay their eggs into plant tissue (White and Elson-Harris 1992). The developing larvae feed on fruit, stalks, leaves, flower heads or seeds, go through three larval instars, and pupate, often in soil prior to emerging as adult flies.

Temperature, humidity and host plants have significant effects on the biology and ecology of fruit flies. Developmental times of embryos and larvae vary at different temperatures, and humidity is necessary for survival – at extreme levels both of these factors can be responsible for increased mortality. It has also been shown that different host plant species can affect rates of development (Bateman 1972).

Pest fruit flies commonly belong to the genera *Anastrepha*, *Bactrocera*, *Ceratitis*, *Dacus* and *Rhagoletis* (White and Elson-Harris 1992). Some species are univoltine (e.g. *Rhagoletis*) while others are multivoltine species (eg. *Bactrocera*), with the latter usually favouring warmer climates, without experiencing winter diapause (Bateman 1972). In Australia, the subfamily Dacinae, including *Bactrocera* and *Dacus,* are highly diverse: over 80 species of *Bactrocera* are endemic, notably the polyphagous pest species *Bactrocera tryoni*, *Bactrocera neohumeralis*, and *Bactrocera jarvisi*.

### 1.1.1 The Queensland fruit fly, *Bactrocera tryoni*

*Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae), commonly known as Queensland fruit fly (Qfly), is a horticultural pest native to Australia. Adults are approximately 7mm long, reddish brown in colour with bright yellow humeral calli (Figure 1.1A and 1.1B). Originally a tropical (and subtropical) rainforest species, *B. tryoni* now has over 40 families of host plants (Hancock *et al.* 2000), many are grown commercially and so the wide distribution of this fly is due, in part, to the expanding availability of host plants as new areas of horticulture were developed. The endemic Australian distribution of *B. tryoni* now covers the east coast of Australia as far south as Gippsland, Victoria, as well as the Northern Territory's tropical north and inland Alice Springs (Figure 1.1C), and is testament to its wide bioclimatic potential (Meats 1981). The distribution has been fairly stable over the last 50 years, with intermittent outbreaks in marginal areas (Dominiak and Daniels 2012, Yonow and Sutherst 1998).

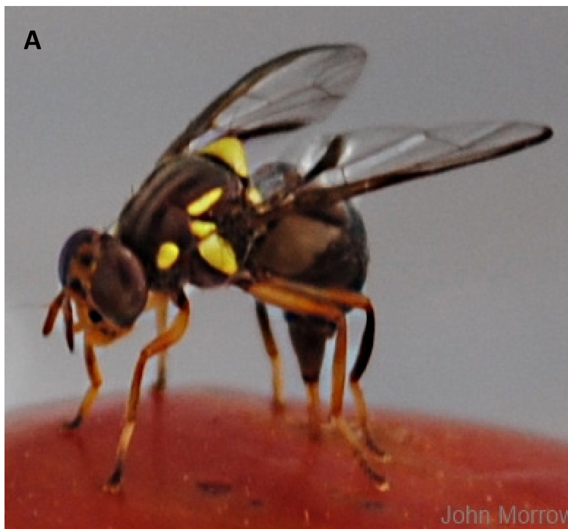


**Figure 1.1** Morphology and distribution of the major horticultural pest species, *Bactrocera tryoni*. **(A)** Female *B. tryoni* ovipositing into a tomato. **(B)** Male *B. tryoni*. (**C)** Current Australian distribution of *B. tryoni* shown in red (Dominiak and Daniels, 2012).

Photos: John Morrow.

Mating activity in *B. tryoni* occurs in the dimming light of dusk (crepuscular) (Tychsen and Fletcher 1971), and the females store the sperm to fertilise eggs as they are laid in clutches of 3-5 under the skin of ripe or ripening fruit (Fitt 1990). The embryos develop over 48h to hatch as first instar larvae (Anderson 1962). The larvae progress through three instars within the fruit, consuming the pulp, which may take 7-10 days, depending on factors including temperature, the type of fruit and the rate of decay of the fruit (Bateman 1972, Christenson and Foote 1960), until exiting the fruit to pupate in the soil over 12 days (Meats 1981). The emerging adult crawls upwards out of the soil, and is ready to mate within one week (PHA 2011). Food sources for adult flies include insect honeydew, fruit, and bird faeces (Bateman 1972), however sexual maturation, particularly development of the ovaries, depends upon consumption of protein which may also derive from bacteria on leaf and fruit surfaces (Drew *et al.* 1983), although this may be insufficient as the primary source of protein (Meats *et al.* 2009). Male *B. tryoni* are attracted to cue lure (Drew and Hooper 1981), which is used for monitoring and control throughout its endemic and non-endemic range.

### 1.1.2 The lesser Queensland fruit fly, *Bactrocera neohumeralis*

*Bactrocera neohumeralis* (Hardy) is a sibling species of *B. tryoni*, and part of the *B. tryoni* species complex. Morphologically, *B. neohumeralis* has a slightly darker body colour than *B. tryoni* and has brown humeral calli (Figure 1.2A), which are yellow in *B. tryoni*. Unlike the dusk mating behaviour of *B. tryoni, B. neohumeralis* mates during the middle of the day (diurnal), but fertile hybrids are readily produced under laboratory conditions with mating around dusk as the dominant behaviour in the F1 generation (Smith 1979). While *B. neohumeralis* has almost as many host plants (Hancock *et al.* 2000), its geographic range is sympatric with *B. tryoni*, but confined to the more coastal regions of Queensland and northern New South Wales (Figure 1.2B). The genetic basis for the different bioclimatic potential of *B. tryoni* and *B. neohumeralis* has long been sought (Meats 2006). Despite the clear mating time dimorphism (Meats *et al.* 2003), and microsatellite analysis showing genetic distinction between the sibling species over their sympatric range (Wang *et al.* 2003),

the genetic similarity in mitochondrial and protein coding nuclear genes suggests that there is still some gene flow occurring between the two species (Morrow *et al.* 2000).

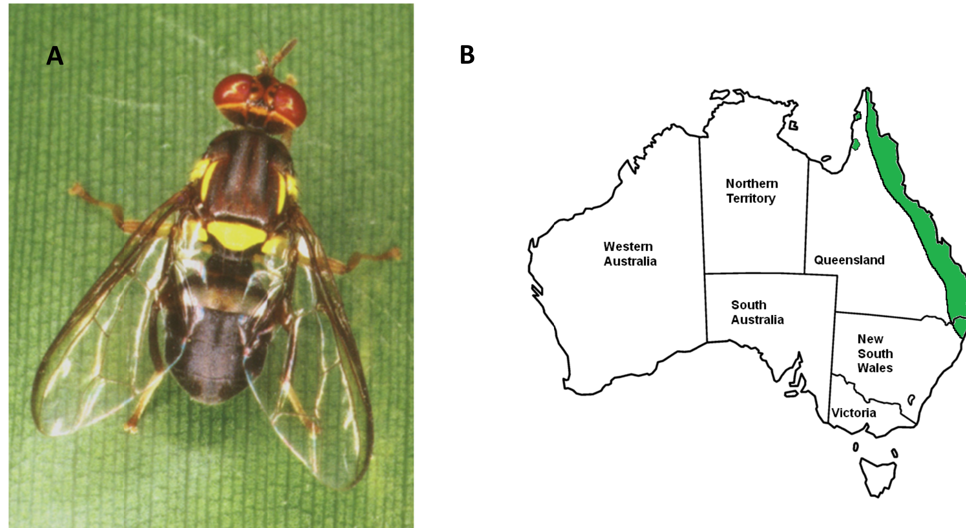



**Figure 1.2** Morphology and distribution of *Bactrocera neohumeralis*. **(A)** Male *B. neohumeralis,* with brown humeral calli (shoulder patches) compared to yellow in *B. tryoni.* *(***B)** Current Australian distribution of *B. neohumeralis.* Photo: Paul Zbrowski.

### 1.1.3 Jarvis's fruit fly, *Bactrocera jarvisi*

*Bactrocera jarvisi* (Tryon) is endemic to Australia, and a pest species in Queensland and the Northern Territory. Its main native host is cocky apple, *Planchonia careya*, and its distribution mirrors that of its main host (Figure 1.3B), although more recently it has been collected from 29 host plant families (Hancock *et al.* 2000), including commercial hosts such as mango, peach, guava and banana (Drew 1989). This species is of medium size, like *B. tryoni* and *B. neohumeralis*, but is golden brown in colour (Figure 1.3A) and is not strongly attracted to cue-lure (Royer and Hancock 2012); an alternative male lure, zingerone, has been developed to attract and monitor *B. jarvisi* distribution (Fay 2012). Fertile hybrids can be created in the laboratory with *B. tryoni* (Cruickshank *et al.* 2001) and *B. neohumeralis* (A. S. Gilchrist, pers. comm.). Drew (1989) placed *B. jarvisi* in the subgenus *Afrodacus*, however this classification has been revised to place it in the same subgenus *Bactrocera* as *B. tryoni* and *B. neohumeralis*. Nevertheless, this species has been useful as an outgroup for molecular studies comparing *B. tryoni* and *B. neohumeralis*

(Morrow *et al.* 2000), as well as a source of mitochondrial and Y-chromosome markers (Shearman *et al.* 2010).

**Figure 1.3** Morphology and distribution of *Bactrocera jarvisi*. **(A)** Female *B. jarvisi*. **(B)** Current Australian distribution of *B. jarvisi.*

Photo: John Morrow

## 1.2 Fruit fly pest control strategies

### 1.2.1 Pesticides and the sterile insect technique (SIT)

*Bactrocera tryoni* is highly invasive, and the historic geographical expansion of *B. tryoni* has coincided with the increase in cultivated fruit beginning in the 1850s (May 1961). The cost of *B. tryoni* to Australian horticulture is measured in damage to fruit and vegetables, disinfestation treatments, restriction of market access and preventative measures. Losses have been estimated at $28.5 million annually (Sutherst *et al.* 2000), with expenditure on fruit fly management estimated at $128 million from 2003 to 2008 (PHA 2008). The restricted movement of potentially infested fruit within Australia and to other parts of the world is integral to containment. Monitoring of the distribution of *B. tryoni* (and related species) in endemic and non-endemic regions by trapping grids utilises the male-attractant

pheromone cue lure, which also attracts a number of other *Bactrocera* species (Osborne *et al.* 1997, Royer and Hancock 2012).

Conventional control strategies have relied heavily on chemical spray applications and lure-and-kill strategies in the field accompanied by post-harvest treatment as a means of disinfestation. Lure-and-kill methods can use food sources (i.e. protein bait) or pheromones (for example cue lure) or coloured sticky traps to either reduce fly numbers or for monitoring outbreaks and incursions. More distantly related species have other male-attractants such as methyl eugenol (Drew and Hooper 1981) and zingerone (Fay 2012).

The recent review and withdrawal of a number of effective organophosphate chemicals has resulted in an acute need for effective alternative control strategies (Dominiak and Ekman 2013). The sterile insect technique (SIT; Klassen and Curtis 2005) is one method that can be an efficient and pesticide-free way to eradicate pest insects, in particular from infested marginal areas where they are not yet established at high population numbers. SIT has already been used effectively to eliminate the New World screwworm from a serious infestation in Libya (Lindquist *et al.* 1992) and is currently exploited to control tephritid fruit flies such as the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae), in a number of world regions including Australia (Robinson 2002).

So far, to control *B. tryoni*, SIT has been employed at a relatively small scale in Australia, in an area that covers temperate fruit growing regions of the states New South Wales, Victoria and South Australia, formerly designated the Fruit Fly Exclusion Zone (FFEZ), and also to control isolated outbreaks in South Australia and Western Australia. Besides this, SIT may also have potential in areas to which *B. tryoni* is endemic. SIT involves the repeated release of vast numbers of irradiated sterile *B. tryoni* in an infested area. Released sterile males compete with field males to mate with field females. Any wild females that mate with a sterile male produce no progeny, and this can reduce or eliminate field populations over successive releases of sterilised flies. At the moment there is no adequate sexing technique for *B. tryoni*, so mixed-sex releases of sterile flies are carried out, potentially reducing the success of this method in population suppression. Improvements to current pest management strategies, as well as new strategies, are vital to better control pests of

agricultural and medical importance. Besides the endemic pest fruit flies, invasive exotic fruit flies that also present a particular biosecurity risk to Australia include *C. capitata*, currently in Western Australia, and the *Bactrocera dorsalis* group of tephritids endemic to south-east Asia (Dominiak and Daniels 2012).

### 1.2.2 Improvements in SIT and mass-rearing techniques

Since 1995, *C. capitata* SIT has benefited from genetic sexing strains (GSS) in mass-rearing facilities, which allow for the exclusive release of male flies to increase efficacy by 3 to 5 times (Robinson 2002). Both first generation GSS based on pupal colour and second generation *temperature sensitive lethal* (*tsl*) demonstrated, in field cage trials, the advantages of male-only release on sterile male dispersal and competitiveness, and the absence of sterile female stings (Hendrichs *et al.* 1995). Implementation of SIT against *B. tryoni* in Australia with a mixed-sex cohort suffers the economic costs of rearing females and the damage caused by sterile female stings, but the improvements that a male-only strain delivers to sterile male dispersal and competitiveness must be trialled, as dispersal behaviour in *B. tryoni* differs from *C. capitata* (Meats and Edgerton 2008).

Tackling the deficiency of a GSS for *B. tryoni* is a key benefit of investigating genes in the sex-determination pathway that are effective targets for RNAi-based transgenic male-only strains (Schetelig *et al.* 2012). Other suggested approaches have exploited tetracycline-repressible expression systems to over-express toxic or pro-apoptotic products (Heinrich and Scott 2000, Thomas *et al.* 2000). An improved protocol, termed the Tet-off transgenic embryonic sexing system (TESS; Schetelig and Handler 2012a), tested in *Anastrepha suspensa* (Diptera: Tephritidae), utilised a genetic construct that may be applicable across a broad range of species, including *Bactrocera* species, if species-specific promoters are found and included in the transgene. The potential for the application of transgenic techniques has been demonstrated for *B. tryoni* (Raphael *et al.* 2004, Raphael *et al.* 2011).

An ancillary problem of mass-rearing insects for SIT is the reduced fitness, both as a result of laboratory adaptation and loss of genetic diversity from inbreeding (Gilchrist *et al.* 2012), and the gamma-radiation sterilising procedure (Parker and

Mehta 2007, Rull *et al.* 2012). The surveys of microbial composition of *Drosophila* spp. indicate that laboratory-reared insects harbour reduced, streamlined or different microbial diversity (Chandler *et al.* 2011) and irradiation distorts the microbial composition of the insect gut, which may contribute to the lower competitiveness (Ben Ami *et al.* 2010). Some of these loss-of-fitness issues can be addressed. Transgenic approaches to genetic sterility can avoid the detrimental fitness effects of irradiation, by producing embryonic lethality in the progeny. The "release of insects carrying a dominant lethal" (RIDL) method, is an alternative that is lethal to female offspring, thereby producing male-only offspring for release. The released males, ideally carrying multiple dominant female-lethal constructs on several chromosomes, produce viable male-only offspring when mated with wild females, which are available to perpetuate this male-only transgenic line in the field until suppression or eradication is achieved (Thomas *et al.* 2000). Alternatively, the conditional expression of lethal pro-apoptotic genes in the progeny of wild females and transgenic males can achieve complete male and female embryonic lethality (Schetelig *et al.* 2009). Similarly, *Wolbachia* can induce cytoplasmic incompatibility that results in embryonic mortality without the need of irradiation (Zabalou *et al.* 2009).

The future of an existing SIT mass-rearing facility for *B. tryoni* in Camden, New South Wales, is currently unclear due to withdrawal of financial contributions deriving from a previous tri-state agreement between New South Wales, Victoria and South Australia. The South Australian Government has recently announced plans to allocate funds towards a similar facility to be built in Port Augusta, South Australia (Warren 2013), recognising the importance of keeping fruit fly out of South Australia and other fruit-growing areas, where there is currently no endemic population. This facility aims to utilise a male-only strain that will be developed through a recently established R&D consortium to optimise the success and efficiency of *B. tryoni* SIT in Australia.

### 1.2.3 Incompatible insect technique (IIT)

A method analogous to SIT that is garnering interest is incompatible insect technique (IIT; Blümel and Russ 1989, Boller *et al.* 1976). This strategy can also result in

suppression of populations, through the mechanism of cytoplasmic incompatibility (CI) between released male insects harbouring strains of the maternally-inherited endosymbiotic bacterium *Wolbachia* that are absent (or different) in field populations. Crosses between *Wolbachia*-infected males and uninfected (or differently infected) females are incompatible and do not yield any offspring. However, females carrying the same (or a compatible) *Wolbachia* strain are fully fertile, impart *Wolbachia* to their progeny and must not, therefore, be released in an IIT approach (Zabalou *et al.* 2009). Thus, while SIT may benefit from releasing male-only cohorts (Hendrichs *et al.* 1995), IIT requires the release of only males.

Furthermore, while *Wolbachia*-based IIT avoids the need for irradiation by inducing crossing sterility, combining IIT and SIT may improve the method, by ensuring that any accidentally-released females are sterilised by irradiation (Zabalou *et al.* 2009). The lower dose required to sterilise females will be less damaging to the fitness of the males, thus maintaining their competitiveness in the field (Arunachalam and Curtis 1985, Brelsfoard *et al.* 2009).

## 1.3 Symbiotic bacteria in insects

### 1.3.1 The biology of *Wolbachia*

It is estimated that 40-65% of insects are infected by *Wolbachia pipientis* (Hilgenboecker *et al.* 2008, Zug and Hammerstein 2012), the only recognised species within this genus of Alphaproteobacteria (Dumler *et al.* 2001, Lo *et al.* 2007), and thus often referred to just by its genus name. *Wolbachia* form an intracellular, symbiotic interaction with the host and are widespread in insects, other arthropods, as well as filarial nematodes (Werren 1997b). In arthropods, *Wolbachia* were first studied in the mosquito *Culex pipiens* (Linnaeus; Diptera: Culicidae) from which the species name, *W. pipientis* was derived (Hertig and Wolbach 1924). Since then, various strains of *Wolbachia* have been identified in many different hosts. Some hosts harbour more than one strain of *Wolbachia* (Jamnongluk *et al.* 2002, Riegler and Stauffer 2002), and these strains have been separated into supergroups based on sequence data, primarily from 16S ribosomal RNA genes (O'Neill *et al.* 1992), the

cell division *ftsZ* gene (Werren *et al.* 1995b) and *Wolbachia* surface protein (*wsp*) DNA (Zhou *et al.* 1998). In more recent years, Multi Locus Sequence Typing (MLST) approaches have been developed based on a number of conserved marker genes (Baldo *et al.* 2006b, Paraskevopoulos *et al.* 2006) or more diverse tandem repeat loci (Riegler *et al.* 2012).

*Wolbachia* are primarily, but not exclusively, localised in gonadal tissue (Dobson *et al.* 1999, Frydman *et al.* 2006) and are transmitted to progeny through maternal deposition in the egg cytoplasm. Most recently, researchers have revealed fitness benefits for *Wolbachia*-infected individuals, for example, in terms of protection from viruses (Hedges *et al.* 2008, Teixeira *et al.* 2008), refractoriness to protists (Moreira *et al.* 2009), improved iron metabolism (Brownlie *et al.* 2009, Hosokawa *et al.* 2010), increased fecundity (Fast *et al.* 2011, Weeks *et al.* 2007) and support of host plant use (Kaiser *et al.* 2010). There are also negative fitness effects in terms of reduced fecundity (Weeks *et al.* 2007), increase cold susceptibility (Maes *et al.* 2012) and desiccation susceptibility (McMeniman and O'Neill 2010). The refractoriness of *Wolbachia*-infected mosquitoes to the Dengue fever virus was successfully tested in cage populations (Walker *et al.* 2011) and *Wolbachia*-induced resistance of the *Aedes aegypti* mosquito to the Dengue virus is currently being trialled in Northern Australia (Hoffmann *et al.* 2011).

However, the significance of *Wolbachia* in this study relates to the potential exploitation, for insect pest management and disease control, of the interesting reproductive phenotypes that are displayed by the hosts in response to their interaction with this symbiont (Figure 1.4). Of specific interest is cytoplasmic incompatibility (CI); other reproductive phenotypes include parthenogenesis, male-killing and feminisation (Werren *et al.* 2008).

### 1.3.2 *Wolbachia* and cytoplasmic incompatibility (CI)

The most widespread phenotype of reproductive manipulation by *Wolbachia* is cytoplasmic incompatibility, a means by which the infected host female obtains a reproductive advantage over uninfected females (Werren 1997b). CI occurs when a

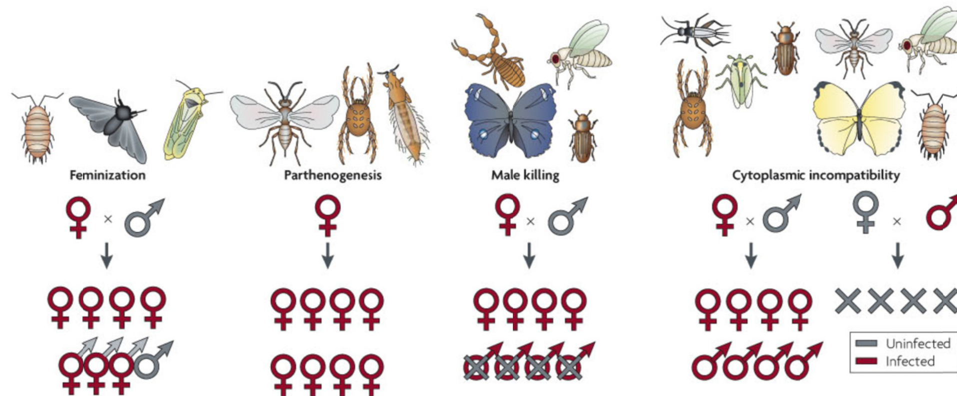



**Figure 1.4** Diagram from Werren et al. (2008) showing the four main reproductive phenotypes caused by *Wolbachia* in different arthropods.

mating between an uninfected female and an infected male results in embryonic mortality, while the *Wolbachia*-infected female can reproduce with infected or uninfected males. As the progeny of an infected female will also carry *Wolbachia* (Werren *et al.* 2008), CI will lead to an increase in infection frequencies in populations over generations, subject to maternal transmission efficiency, level of CI and any other *Wolbachia*-induced fitness effects (Weeks *et al.* 2007).

*Wolbachia* strains can be transferred between host species through embryonic microinjections. In this way, *w*Cer2 and *w*Cer4 from *Rhagoletis cerasi* (Linnaeus; Diptera: Tephritidae) were successfully transferred to *C. capitata* (Zabalou *et al.* 2004) and *w*Cer2 was also established in *Bactrocera oleae* (Rossi) (Diptera: Tephritidae) following transinfection from the *C. capitata* donor (Apostolaki *et al.* 2011). In both of these new hosts, maternal transmission of *Wolbachia* to progeny and CI were complete (100%). However, this is not true for every *Wolbachia*-host interaction; not all expressions of CI are complete and transmission to offspring may not be 100%. The CI phenotype is subject to many factors, including host species, *Wolbachia* genotype, *Wolbachia*-host genotype interactions as well as environmental factors (Werren *et al.* 2008).

Although the molecular mechanism underlying CI is still unclear (Pinto *et al.* 2013, Serbus *et al.* 2008), the cytological mechanism by which CI causes embryonic mortality has been detailed in several hosts. CI acts on the crucial synchronised alignment of chromosomes during the first mitotic division. As described by the

current model, a sperm modification is made during spermatogenesis so that the breakdown of the nuclear envelope after egg fertilisation is delayed (Landmann *et al.* 2009). Consequently, paternal chromosomes are out of phase with the maternal chromosomes in zygotes derived from incompatible matings (Tram and Sullivan 2002). This results in a haploid embryo which is not viable in host species that require both chromosome sets for development, including *Drosophila* and fruit flies in general.

The modification to the sperm nuclei causing incompatibility is rescued by the presence of the same strain of *Wolbachia* in the egg, through transmission from an infected mother. *Wolbachia* present in the egg realigns the timing of the mitotic phase of the zygote (Tram and Sullivan 2002), in part by *Wolbachia*-mediated up-regulation of host genes (Pinto *et al.* 2013). This binary mechanism is referred to as modification (mod) / rescue (res) (Werren 1997a). Sometimes mod / res can also be successful between different strains of *Wolbachia* (Charlat *et al.* 2004). Alternatively such interactions can be partial, for example, one of the strains lacks the modification ability, but it can still rescue the modifications caused by closely related strains (Bourtzis *et al.* 1998), and there are data to suggest that some strains may hold the means to rescue more than one CI sperm modification (Zabalou *et al.* 2008).

### 1.3.3 *Wolbachia* and parthenogenesis, male-killing and feminisation

Parthenogenesis is a mode of reproduction where development of embryos proceeds in the absence of fertilisation. *Wolbachia* have been shown to induce parthenogenesis in insect groups that have a haplo (male) – diploid (female) sex-determination system, for example, wasps (Stouthamer *et al.* 1990), mites (Weeks and Breeuwer 2001) and thrips (Arakaki *et al.* 2001), all of which can produce male progeny from unfertilised eggs independent of *Wolbachia* (=arrhenotoky). The effect of *Wolbachia* is to cause unfertilised eggs to develop into females (=thelytoky) (Stouthamer *et al.* 1990).

*Wolbachia*-induced male-killing has been described in a range of species within the orders of Coleoptera, Lepidoptera and Diptera. In the adzuki bean borer *Ostrinia scapulalis* (Walker) (Lepidoptera: Crambidae), an insect that employs a ZW/ZZ sex-

determination system with heterogametic females, *Wolbachia* will kill chromosomal males (ZZ) during larval development. In lines cured of *Wolbachia*, genetic females (ZW) die during larval development, whereas males are viable. This research indicated that female development has become dependent on *Wolbachia*, and that *Wolbachia* is detrimental to development of ZZ males (Kageyama and Traut 2004).

Feminisation by *Wolbachia* has been described in woodlice (Crustacea: Isopoda) in terms of a hormonal inhibition affecting the androgenic gland, resulting in chromosomal males (ZZ) developing into females (Vandekerckhove *et al.* 2003). A different mechanism is used in some insects, possibly related to sex determination, as demonstrated in the butterfly *Eurema hecabe* (Linnaeus) (Lepidoptera: Pieridae) (Narita *et al.* 2007) and leafhopper *Zyginidia pullula* (Boheman) (Hemiptera: Cicadellidae) (Negri *et al.* 2006). In *E. hacabe*, the sustained presence of *Wolbachia* throughout larval development is required to maintain the feminising process, whereas antibiotic treatment administered at various stages of development resulted in intersexes (Narita *et al.* 2007). In *Z. pullula,* an XX/X0 sex-determination system has also given rise to sexual mosaicism with *Wolbachia*-induced feminisation. Sexual mosaics are a potential outcome of cell-autonomous somatic sex determination, as sometimes reported in transient RNA interference of key genes in the pathway. In *Z. pullula* the degree of feminisation is dependent on *Wolbachia* density, and may be driven by the manipulation of DNA methylation to follow the female pattern (Negri *et al.* 2009).

### 1.3.4 *Wolbachia* and effects on the sex-determination gene *Sex-lethal*

*Wolbachia* are able to manipulate the reproductive phenotype of their hosts. Remarkably, the affected hosts encompass a range of sex-determination systems, including XX/XY *Drosophila* systems, ZW/ZZ system found in many lepidopterans, XX/XO system found in a hemipteran and haplo-diploidy in parasitoids and thrips. Some of these phenotypes hint at possible *Wolbachia* interactions with sex-determination genes in these organisms. For example, the male-killing phenotype induced by another bacterial symbiont, *Spiroplasma poulsonii* in *Drosophila melanogaster* Meigen (Diptera: Drosophilidae), requires the formation of a functional dosage compensation complex (Veneti *et al.* 2005). Conversely, there is

also evidence for sex-specific regulation of *Wolbachia* genes in *Culex* mosquitoes, indicating that *Wolbachia* can adapt sex-specific interactions with hosts (Papafotiou *et al.* 2011, Walker *et al.* 2007).

In *D. melanogaster*, where *Sxl* is the master gene, it seems that *Wolbachia* may have an effect on the role of *Sxl* in oogenesis, but not in somatic sex determination (Sun and Cline 2009). *Wolbachia* rescues certain *D. melanogaster Sxl* mutant lines from defective egg production. Other mutant lines with similar phenotypes are not rescued by *Wolbachia*, therefore a particular interaction between *Sxl* and *Wolbachia* is suspected (Starr and Cline 2002).

### 1.3.5 Other symbiotic microorganisms in tephritid fruit fly species

Symbiotic microorganisms other than *Wolbachia* are an important and common part of insect life, with effects on their hosts ranging from harmful to beneficial. Examples of mutualistic symbioses are found in insects that rely on gut microorganisms to provide nutritional supplements when their diet is primarily plant material, including feeders of plant sap (Fukatsu and Hosokawa 2002), wood (Hongoh *et al.* 2008, Warnecke *et al.* 2007) and fruit (Petri 1910). Other functions include enhanced adaptability to food sources (Tsuchida *et al.* 2009), nutrition (Warnecke *et al.* 2007), increased immunity (reviewed in Broderick and Lemaitre 2012), protection from parasites (Koch and Schmid-Hempel 2011), and toxin degradation (Kikuchi *et al.* 2012). Symbiotic microorganisms can also influence mating preference (Sharon *et al.* 2010) and have detrimental effects such as increased susceptibility to toxins (Broderick *et al.* 2006) and attraction of predators (Leroy *et al.* 2011).

Evidence for definitive functional roles of symbionts in tephritid fruit flies is scarce – *B. tryoni* may benefit from nitrogen fixation (Murphy *et al.* 1994), *C. capitata* appears to gain nitrogen fixing and pectinolase functionality (Behar *et al.* 2005), but *B. oleae* is, so far, the only species with a strong candidate for an obligate symbiont (Capuzzo *et al.* 2005, Petri 1910), although this bacterium, *Candidatus* Erwinia dacicola, was not detected in *B. oleae* laboratory lines (Kounatidis *et al.* 2009). The presence of highly-similar bacteria in related species of the tephritid subfamily

Tephritinae, which inhabits Asteraceae flower-heads suggests co-evolutionary interactions (Mazzon *et al.* 2008). In tephritid fruit flies, symbionts may have implications for host plant preference, adaptability and field fitness which can potentially be exploited or manipulated for control of pests (Jurkevitch 2011).

Vertical transmission of microbes can be achieved transovarially, such as for *Wolbachia*; in other cases, morphological adaptations are made, to ensure the transmission of microbes to the next generation (Bateman 1972). Many tephritid fruit flies possess an oesophageal bulb (or diverticulum), which in the olive fly *B. oleae* almost exclusively harbours *Candidatus* E. dacicola (Capuzzo *et al.* 2005). This obligate symbiont is released into the midgut and is also found on the ovipositor and in unhatched larvae prior to feeding, indicative of maternal transmission through smearing of the egg during oviposition (Capuzzo *et al.* 2005, Estes *et al.* 2009, Petri 1910). Plataspid stinkbugs provide important bacteria to their offspring by ingestion of symbiont capsules laid beneath the egg mass (Fukatsu and Hosokawa 2002). Insects also acquire symbionts from food, social interactions and the environment (reviewed in Engel and Moran 2013).

The concept that a core microbiome is present has not been shown in *Drosophila* species, where the microbial community has been described as both inconstant and low in diversity (Wong *et al.* 2011, Wong *et al.* 2013); this also applies to the majority of tephritids, where various culture-dependent and -independent studies identified variability in composition and abundance at genus and species taxonomic levels, but an overall consistent dominance of few bacterial families (reviewed in Behar *et al.* 2009). The inconstancy of the microbiome may be predicated on the instability of the insect gut, which sheds its lining at each larval moult and only becomes stable at the adult stage (Engel and Moran 2013).

### 1.4 *Drosophila melanogaster* somatic sex determination pathway

Current understanding of the sex determination pathway of tephritid fruit flies is best examined in light of the much more comprehensively described pathway in

*D. melanogaster*, although there are a number of significant differences between the sex-determination pathway between *Drosophila* and tephritids.

### 1.4.1 The primary signal in *D. melanogaster*

The primary signal for sex determination in *D. melanogaster* is the X-chromosome dose (Erickson and Quintero 2007) – two X chromosomes in females and one X-chromosome in males. In this species, the Y-chromosome does not play a role in sex determination but is necessary for male fertility. The primary signal is communicated to the key gene, *Sex-lethal* (*Sxl*) by the products of the X-linked signal element (XSE) genes when combined with maternally-derived and autosomal-linked protein co-factors (Schutt and Nothiger 2000). It is this complex interaction between the dosage of many gene products that initiates the production of functional SXL through activation of the establishment promoter (*Sxl*Pe) (Schutt and Nothiger 2000).

Four XSE genes have been identified – *sisterless A* (*sisA*), *scute* (*sc*), *outstretched* (*os*; formerly *unpaired*) and *runt* (*run*). These genes are transcription factors that have other roles in development, but appear to have been co-opted as XSEs to target *Sxl*.

- *scute* is also a proneural gene (Torres and Sanchez 1989), but during early development it forms a heterodimer with the maternally-derived protein *daughterless* (DA) to bind directly to *Sxl*Pe (Yang *et al.* 2001).
- *runt* is involved in segmentation and belongs to a class of DNA binding proteins (Kramer *et al.* 1999, Torres and Sanchez 1992).
- *sisA* has a role in posterior- and anterior-midgut and endoderm development, and encodes a basic leucine zipper domain (bZIP) protein that activates transcription of *Sxl* (Erickson and Cline 1993). Both *sisA* and *sc* belong to a family of basic helix-loop-helix (bHLH) transcription activators.
- *os* triggers JAK kinase to activate the STAT92E transcription factor (Harrison *et al.* 1998). This process occurs during the 13$^{th}$ cycle of nuclear division (Avila and Erickson 2007), much later than the onset of *sisA* in cycle 8, *sc* in cycle 9, and *Sxl*Pe in cycle 12 (Erickson and Cline 1993). Detection of *os* expression after *Sxl*

promotion was attributed to its role in reinforcing the *Sxl*Pe activation initiated by *sisA*, *sc* and *run* (Avila and Erickson 2007).

Maternal products, in addition to *da*, include *hermaphrodite* (*her*), *extra-macrochaetae* (*emc*) and *groucho (gro)*. DA has many roles in addition to positive regulation of sex determination, and is expressed throughout development in both sexes (Cronmiller *et al.* 1988). The initiation of *Sxl*Pe requires *her* protein (HER), but its zygotic expression is also necessary for normal female development (Pultz and Baker 1995).

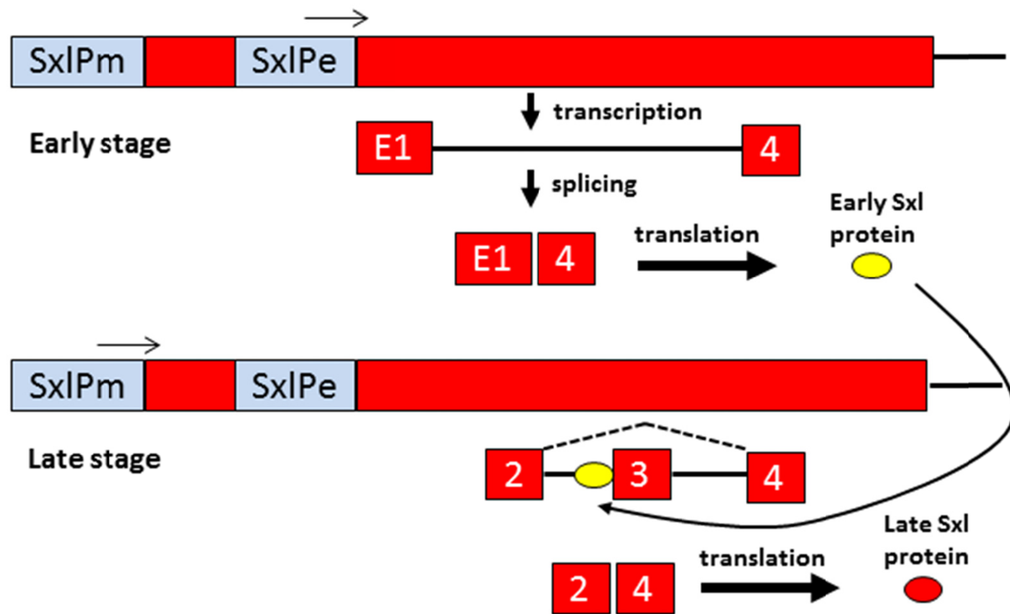
Negative regulators of *Sxl*Pe are *emc* and *gro* and the autosomal gene *deadpan* (*dpn*). The protein encoded by *emc* can bind to bHLH proteins like SC and SISA, and thus prevent binding to the *Sxl*Pe, which reduces the level of *Sxl* activation (Campuzano 2001). GRO is a transcriptional co-repressor and, when interacting with DPN (a DNA-binding protein), an operational unit is formed, further reducing the effectiveness of the positive regulators (Barbash and Cline 1995, Paroush *et al.* 1994).

The additive effects of the maternal products, autosomal gene *dpn*, and zygotically expressed XSEs results in activation of *Sxl*Pe only in females. The outcome is a functional female protein SXL, which is zygotically expressed well before embryonic transition to cellular blastoderm, but no functioning product in males.

### 1.4.2 The key gene in *D. melanogaster* – *Sex-lethal*

*Sex-lethal* is the key gene, because it establishes the determined state of sexual differentiation of somatic tissue in *D. melanogaster* (Cline 1978). *Sxl* encodes an RNA-binding protein that is under the separate control of an early (establishment – *Sxl*Pe) and a late (maintenance – *Sxl*Pm) promoter. *Sxl*Pe supports the transcription of pre-mRNA that is spliced such that a threshold level of functional protein is created only in females (Figure 1.5). In response to the XSE primary signal, the *Sxl*Pe is transiently active – from nuclear cycle 12 to 14 – and must generate sufficient SXL to prime the next reaction. The *Sxl*Pm is activated in both males and females; however, this occurs earlier in females and bolsters the chances of

**A. Female embryo**




**B. Male embryo**

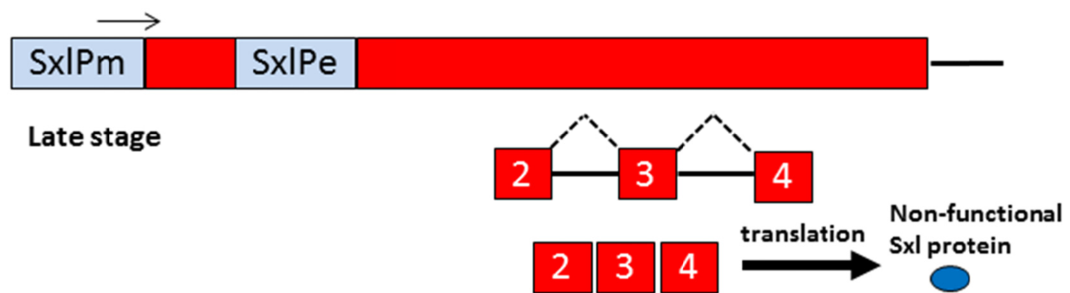



**Figure 1.5** Expression of *D. melanogaster Sxl* gene from the early (*Sxl*Pe) and late (*Sxl*Pm) promoters; only the first four exons are shown. **(A)** In females, transcription is first initiated from *Sxl*Pe in response to the primary signal; the early SXL protein becomes part of the regulation of splicing of transcripts from the *Sxl*Pm, leading to continuous production of functional SXL. **(B)** In males, insufficient or no early SXL is made; in its absence exon 3 is included in the mature transcript, which encodes a stop codon and results in a non-functional protein. Diagram adapted from Molecular Biology Web Book, by Frank Lee.

successful establishment of the female sex-determination pathway. In females, *Sxl*Pm is activated in cycle 13, before *Sxl*Pe is finished, but in males *Sxl*Pm starts in cycle 14 (González *et al.* 2008). The *Sx*lPm can produce two different transcripts after processing (Figure 1.5). The default mRNA includes exon 3, which introduces a stop codon and yields a truncated protein without function. In the presence of SXL, an alternate splicing pattern is used which excludes exon 3 from the mRNA, so an open reading frame (ORF) is restored and active SXL is generated (Horabin and Schedl 1993, Sakamoto *et al.* 1992).

SXL has a secondary role, binding to the 3' untranslated region of the pre-mRNA and suppressing translation of the mature *Sxl* transcript. This negative feedback loop in conjunction with the positive feedback loop allows for a constant amount of SXL to be produced, avoiding over-expression causing toxicity (Yanowitz *et al.* 1999).

### 1.4.3 Signal transduction genes – *transformer* and *transformer-2*

*Sex-lethal* communicates the female-determined state through controlling the splicing of *transformer* (*tra*). Again, the female state requires the production of an active TRA protein, while in males a non-functional product is generated through default splicing of the *tra* pre-mRNA.

The *tra* gene is located on chromosome 3 and produces two mature transcripts, one found in both males and females, the second found only in females (Boggs *et al.* 1987, Sosnowski *et al.* 1989). The genetic structure contains exons 1, 2 and 3, plus a non-sex-specific exon located between exons 1 and 2, adjacent to the 5' end of exon 2 (Figure 1.6). Both transcripts use the same splicing donor site at the 5' end of intron 1, however, exon 2 is ligated directly to exon 1 in female-specific transcripts, but non-sex-specific transcripts are produced when exon 1 is ligated to the non-sex-specific exon. As a result, 175bp of additional sequence is incorporated into the non-sex-specific transcript and encodes stop codons, which generate truncated proteins. The female-specific mRNA has an open reading frame and produces a functional protein of approximately 22kDa (Boggs *et al.* 1987).

In general, the splicing of mRNA is performed by the spliceosome which binds directly to RNA signal sequences. An essential part of this assembly is U2AF (U2 auxilliary factor). U2AF has very high affinity for a range of polypyrimidine tracts, and its binding efficiency to the *tra* non-sex-specific splice site is 100-fold greater than for the female-specific splice site (Valcarcel *et al.* 1993). Therefore, in male cells, the binding of U2AF to the proximal 3' splice site is highly favoured, and leads to the non-sex-specific transcript (Figure 1.6).

The mechanism responsible for *tra* sex-specific splicing requires the preferential binding of SXL to poly(U) sequences in *tra* pre-mRNA to block the default mode of splicing. The binding efficiency of SXL to the *tra* non-sex-specific site is even

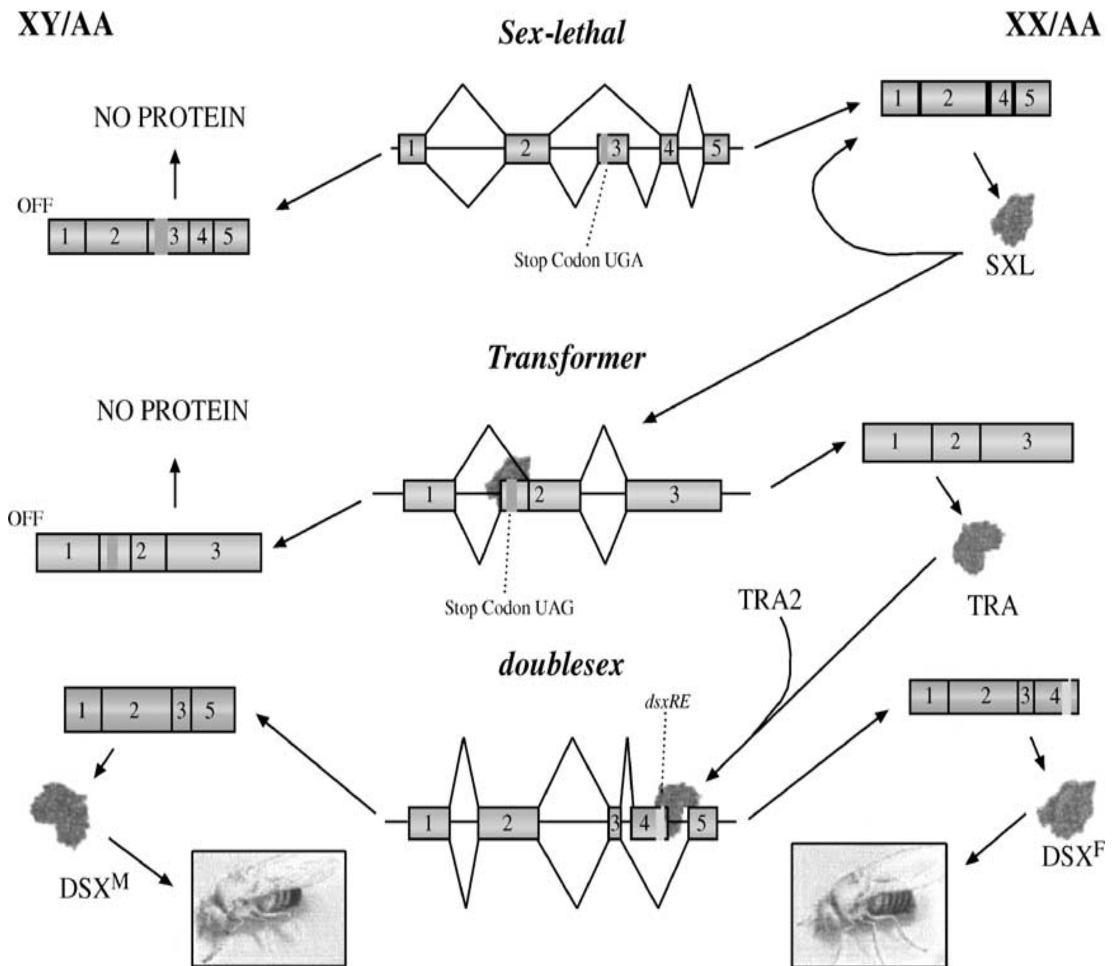



**Figure 1.6** Simplified sex-determination pathway in *D. melanogaster*, where alternative splicing of mRNA leads to different protein products in males and females. In XX females, a functional SXL protein is generated (see Figure 1.5) which regulates the female-specific splicing of the transformer transcript. The functional TRA product, combined with TRA-2 promotes the female-specific splicing of *dsx*. In XY males, the lack of functional SXL and TRA leads to male-specific splicing of *dsx* (from Saccone *et al.* 2002).

greater than that of U2AF. Valcarcel *et al*. (1993) showed the specific binding affinity of SXL to *tra* non-sex-specific poly(U) is 15 to 100 times greater than its affinity for any other polypyrimidine tract, and that its greater affinity for this particular *tra* sequence is sufficient to bind stably in the face of U2AF competition. When displaced, U2AF binds to the less ideal distal 3' site (female-specific) and effects splicing at that location (Valcarcel *et al.* 1993). This blocking mechanism is not complete – both males and females express the non-specific *tra* mRNA. Approximately half of the *tra* pre-mRNA is spliced in the female form due to the action of SXL (Boggs *et al.* 1987), but half follows the default splicing path.

The role of *transformer*-2 (*tra-2*) differs in males and females. The transcript is not spliced in a sex-specific manner in the soma, but shows tissue-specific splicing in males. The *tra-2* gene encodes an RNA-binding protein in both male and female *D. melanogaster* (Amrein *et al.* 1988). Experimentation with *tra-2* mutants shows that TRA-2 is necessary continuously throughout development for female somatic differentiation (Belote and Baker 1982) and for spermatogenesis in males (Schüpbach 1982). The expression level of *tra-2* is significantly higher in females (Amrein *et al.* 1988).

Both TRA and TRA-2 contain regions characteristic of proteins with splicing regulatory function. An arginine/serine-rich domain (RS domain) is found in both the TRA (Boggs *et al.* 1987) and TRA-2 (Goralski *et al.* 1989) protein sequences, and this region has a function in mRNA splicing (Huang and Steitz 2005) through its interaction with proteins and RNA. TRA-2 has an additional RNA binding domain (RBD) of the type RNP-CS, a common eight amino acid motif found in many proteins in many organisms (Bandziulis *et al.* 1989). TRA and TRA-2 form an assemblage with other splicing regulator (SR) proteins, notably RBP-1, a demonstrably important component of this complex for enhancing appropriate spliceosome assembly on *doublesex* (*dsx)* and *fruitless* (*fru)* pre-mRNAs (Heinrichs and Baker 1995, Heinrichs *et al.* 1998).

### 1.4.4 The genetic doubleswitch – *doublesex*

The *dsx* pre-mRNA is spliced into two sex-specific, functional transcripts. In female cells, TRA and TRA-2 form a complex with other SR proteins to induce production of female-specific $DSX^F$ containing three common exons (exons 1-3) and the female-specific exon 4 (Figure 1.6). In male cells deficient in TRA, the default mode produces $DSX^M$ containing three common exons (exons 1-3) and exons 5 and 6 (Burtis and Baker 1989).

Splicing of *dsx* in the common regions and the male-specific regions is due to sequence suitability as splice donor and acceptor sites, according to the splicing acceptor consensus (T/C)nNCAG described in Oshima and Gotoh (1987). However, the female-specific splice acceptor site (upstream of exon 4) is clearly weaker than

its counterpart at male-specific exon 5 (Burtis and Baker 1989). This substantiates the theory that *dsx* is processed in males in the default pathway, and that intervention is required to alter the transcript to the female form (Nagoshi *et al.* 1988).

Further analysis of the female-specific mRNA revealed the presence of six conserved 13-nucleotide sequences, located in the 3' untranslated region of exon 4 (Burtis and Baker 1989) and shown to bind TRA and TRA-2 (Hedley and Maniatis 1991, Inoue *et al.* 1992). These conserved sequences are exonic splicing enhancers (ESEs) and are essential to signal the assembly of TRA and TRA-2 into a complex to promote the use of the weak female-specific acceptor site (Ryner and Baker 1991). The translation start site is located in exon 2 and generates a 427 amino acid product ($DSX^F$) which is responsible for somatic cell differentiation. In collaboration with *her* and *intersex* (*ix*), $DSX^F$ activates genes involved with female differentiation, and suppresses those involved with male somatic development. In male *D. melanogaster*, a functional $DSX^M$ protein of 549 amino acids is produced in the absence of TRA and is similarly responsible for the male differentiation of the soma (Burtis and Baker 1989).

Both forms of DSX are transcription factors and contain a DNA-binding domain in the common region. The region called oligomerisation domain 1 (OD1) is located in positions 39 to 104 of the protein, OD2 begins at 350 and terminates at the stop codon, position 426 in $DSX^F$; and 456 in $DSX^M$. OD1 contains a zinc finger domain and functions in DNA binding and protein oligomerisation, and OD2 extends into the sex-specific domains and is used for protein / protein interactions (An *et al.* 1996, Cho and Wensink 1997). Yolk protein genes appear to be targets of DSX activation (Burtis *et al.* 1991), and other candidates include genes involved in sex comb bristle formation and pigmentation (Jursnich and Burtis 1993).

Two forms of *fru* protein, $FRU^F$ and $FRU^M$ are also generated in the presence ($FRU^F$) or absence ($FRU^M$) of the TRA/TRA2 complex. These transcription factors control the expression of target genes associated with female and male behaviour respectively (Ryner *et al.* 1996).

## 1.5 The somatic sex-determination pathway in Tephritidae

With the serendipitous discovery of a sex chromosomal marker allowing the differentiation of *C. capitata* at early developmental stages (Gabrieli *et al.* 2010), the study of sex-determination genes in *C. capitata*, and also *B. oleae* and *B. tryoni* (Table 1.1), has gained momentum with the potential for a genetic sexing strain to improve SIT efficiency in the field, as well as for general interest in insect sex determination.

| | *B. tryoni* | *B. jarvisi* | *B. oleae* | *C. capitata* | *Anastrepha* spp. | *Drosophila melanogaster* | *Musca domestica* |
|---|---|---|---|---|---|---|---|
| *Sxl* | Chapter 5 | Chapter 5 | Lagos et al. 2005 | Saccone et al. 1998 | | Cline 1978 | Meise et al. 1998 |
| *tra* | Chapter 5 | Chapter 5 | Lagos et al. 2007 | Pane et al. 2002 | Ruiz et al. 2007; Schetelig et al. 2012 | Belote et al. 1989 | Hilfiker-Kleiner et al. 1993 |
| *tra-2* | Chapter 5 | Chapter 5 | | Salvemini et al. 2009 | Sarno et al. 2010; Schetelig et al. 2012 | Belote & Baker, 1982 | Burghardt et al. 2005 |
| *dsx* | Shearman & Frommer, 1998 | | Lagos et al. 2005 | Saccone et al. 2008 | Ruiz et al. 2005 | Burtis & Baker, 1989 | Hediger et al. 2004 |

**Table 1.1** References for the primary sex-determination genes in tephritids, *Drosophila melanogaster* and the house fly, *Musca domestica.*

Research into the sex-determination genes of Tephritidae has thus far revealed the presence of homologues of the fundamental sex-determination genes of *D. melanogaster*. However, the regulation and function of certain genes has diverged. The evolution of regulatory pathways is hypothetically built from an ancestral base, and more recent acquisitions are recruited upstream. "Masters change, slaves remain" (Graham *et al.* 2003) concisely summarises this supposition, and is indeed supported by the comparison of Tephritidae and Drosophilidae. The roles of *doublesex* are strongly conserved, and those of *transformer-2* and *transformer* are somewhat conserved. However, regulation of *tra* is different and, unlike *D. melanogaster*, *Sex-lethal* exhibits no sex-specific expression. Even more striking is that the primary signal is located on the Y-chromosome (Willhoeft and Franz 1996b) and is not dependent on the number of X-chromosomes.

The primary signal of many members of the order Diptera is the *Dominant Male Determiner (M*). The nature of *M* has yet to be ascertained, but its location on the Y-chromosome is demonstrated by the existence of *C. capitata* males with XXY chromosomal complement and XO females (Willhoeft and Franz 1996a, Willhoeft and Franz 1996b). In *C. capitata* and *B. tryoni*, male phenotype is dependent on the presence of the Y-chromosome (Meats *et al.* 2002, Willhoeft and Franz 1996b, Zhao *et al.* 1998). Both X and Y-chromosomes of *B. tryoni* are heterochromatic (Zhao *et al.* 1998) which suggests that a limited number of genes are located on these chromosomes.

The gene that is targeted by *M* is unlikely to be *Sex-lethal* because it is not sex-specifically regulated and an active protein is generated in both sexes in *C. capitata* (Saccone *et al.* 1998) and *B. oleae* (Lagos *et al.* 2005) (Table 1.1). The function of SXL in these fruit flies has yet to be ascertained. The more likely candidate is *tra* mRNA or protein, in its role as the key gene, but it is also possible that transient interference with other splicing factors such as *tra-2* will have the same effect.

### 1.5.1 The key gene – *transformer*

Homologues of the *D. melanogaster* gene *transformer* (*tra*) have been identified in tephritids (*C. capitata, B. oleae*, *Anastrepha* spp.) and in other Diptera such as the housefly *Musca domestica* Linnaeus (Diptera: Muscidae) and sheep blowfly *Lucillia cuprina* (Wiedemann) (Diptera: Calliphoridae) (Table 1.1; Concha and Scott 2009). Sequence and expression analysis clearly illustrates the difference between the splicing pattern in male and female adults. The genetic structure of the *tra* locus is similar within the tephritid family and orthologous to *D. melanogaster tra*. In *B. oleae*, exons 1A and 1B (equivalent to Exon 1 in *C. capitata*), exons 2A and 2B (equivalent to exon 2 in *C. capitata*) and exon 3 are common in all transcripts, and make up the entire female-specific transcript (Lagos *et al.* 2007, Pane *et al.* 2002). The *C. capitata* protein CcTRA exhibits functional conservation with *D. melanogaster*, as demonstrated by its transformation of *D. melanogaster* XY males into phenotypic females through promotion of a heat-shock-coupled, full-length Cc*tra* transgene (Pane *et al.* 2005). RNAi experiments have further confirmed the importance of TRA in *C. capitata* and *B. oleae*, with double-stranded RNAs

containing the *tra* transcript capable of interfering with the sex-determination pathway to yield fertile XX males, which produce only female offspring (Lagos *et al.* 2007, Pane *et al.* 2002).

Like *D. melanogaster*, the inclusion of male-specific sequence in *C. capitata,* encoding stop codons in the region between exons 1 and 2, renders the product non-functional (Pane *et al.* 2002). A similar mechanism is found in *B. oleae* (Lagos *et al.* 2007) and *Anastrepha* species (Ruiz *et al.* 2007, Schetelig *et al.* 2012). However, unlike *Drosophila*, up to four male-specific sequences are spliced into the mature *tra* mRNA of tephritids (Lagos *et al.* 2007, Pane *et al.* 2002, Ruiz *et al.* 2007) suggesting a more complex mechanism of splicing is in action. Furthermore, the identification of eight TRA/TRA2 binding sites (13-nt consensus sequences), as found in *D. melanogaster dsx*, led to the interpretation that *tra* is regulated by an auto-regulatory loop of TRA and TRA-2 proteins (Pane *et al.* 2002), probably without intervention by SXL.

Although some investigators found traces of other partially spliced *tra* mRNA sequences in adult female *C. capitata* (Pane *et al.* 2002), others reported no detectable male transcript after 9h development (Gabrieli *et al.* 2010). *Bactrocera oleae* and *Anastrepha obliqua* (Macquart) (Diptera: Tephritidae) expression analyses reveal no indication of male-specific *tra* mRNA in adult females (Lagos *et al.* 2007, Ruiz *et al.* 2007). This is an interesting divergence from *tra* expression in *D. melanogaster*, where females express both transcripts equally throughout development, with only the female form being productive. This difference may perhaps be attributed to the efficiency of the splicing regulators – in *Drosophila* it is SXL, but in tephritids TRA production is likely to be regulated by a feedback loop featuring TRA and TRA-2.

### 1.5.2 *transformer-2*

Many dipteran species express homologues of *D. melanogaster transformer-2*. The gene contains eight exons in *M. domestica* (Burghardt *et al.* 2005), *C. capitata* (Salvemini *et al.* 2009) and *B. oleae* (GenBank Accession No. AJ715415). Expression analysis has identified a single transcript in both sexes at all stages of

development including unfertilised eggs (Burghardt *et al.* 2005, Salvemini *et al.* 2009). Unlike *D. melanogaster*, no tissue-specific splicing or usage of transcriptional start sites was encountered.

The protein sequence of TRA-2 reveals four conserved features that highlight its function as a splicing regulator. Two RS regions (RS1 and RS2) are found, not highly conserved but replete with arginine and serine residues; a seventy-two amino acid region called RNA-recognition motif (RRM) is highly conserved between *D. melanogaster*, *M. domestica* and *C. capitata*, and an 18 amino acid linker region located between RRM and RS2 is likewise highly conserved (Amrein *et al.* 1988, Burghardt *et al.* 2005, Salvemini *et al.* 2009).

It is probable that TRA-2 forms a complex with TRA as modelled in *D. melanogaster.* The conservation of protein structure suggests a similar function. RNAi experimentation in *C. capitata* and *A. suspensa* substantiates the essential role TRA-2 plays in female development. Significant skewing of the sex ratio toward male phenotype and subsequent karyotype analysis revealed transient expression of *tra-2* was sufficient to alter the natural developmental pathway (Salvemini *et al.* 2009, Schetelig *et al.* 2012). Furthermore, the transcripts detected in adult intersexes and phenotypic males (both chromosomally XX females) revealed that *tra*, *dsx* and *fru* were all affected by the transient loss of *tra-2* (Salvemini *et al.* 2009).

### 1.5.3 *doublesex*

The organisation of *dsx* homologues in tephritid fruit flies, and the functionality of both male and female transcripts, is very similar to that which has been described in *D. melanogaster*. In *B. tryoni*, translation also begins in exon 2 and encodes an ORF to exon 4 in females; while the male mRNA encompasses exons 2 and 3, exon 4 is removed so that exon 3 is ligated directly to exons 5 and 6 (Shearman and Frommer 1998). Again, the existence of 13-nt repeat elements followed by purine-rich sequences in the 3' untranslated region is evidence of a mechanism employing TRA/TRA2 enhancement of alternative splice sites (Shearman and Frommer 1998).

Oligomerisation domains OD1 and OD2 as described in *D. melanogaster* exhibit high homology within fruit fly species and with *D. melanogaster*: *B. tryoni*

(Shearman and Frommer 1998), *B. oleae* (Lagos *et al.* 2005), *C. capitata* (Saccone *et al.* 2008) and *A. obliqua* (Ruiz *et al.* 2005) share homology in OD1 and both the male- and female-specific regions of OD2 of more than 95%. Indeed, OD1 of these tephritids, located in the common region of DSX, is highly similar to that of *D. melanogaster,* with a calculated homology upwards of 97%.

Saccone *et al.* (2008) demonstrated that $CcDSX^M$, expressed in *D. melanogaster*, changed chromosomal females into phenotypic males, thus confirming the functional equivalence of the *dsx* orthologues. One might suppose that the target genes of *dsx* may be similar or identical in the different Dipteran families, and it will be of interest to investigate the level of accord between the downstream target genes in both Tephritidae and Drosophilidae.

The investigation into the sex-determination genes of *B. tryoni* can provide a target to manipulate the genetic pathway and engineer a line of *B. tryoni* that produces exclusively males for release in SIT and IIT. There are several candidate genes for this role, such as the homologues of *D. melanogaster* genes, *transformer* and *transformer-2*, that appear to coordinate the sex-specific splicing of both *transformer* itself, and genes further down the sex-determination pathway in non-drosophilid dipterans (Lagos *et al.* 2007, Pane *et al.* 2002, Salvemini *et al.* 2009, Schetelig *et al.* 2012), or *M*, the as-yet-to-be-identified agent that is expected to interfere with either *transformer* or *transformer-2* to steer development away from the female pathway and down the male pathway (Shearman 2002).

## 1.6 Research scope and aims

The research detailed in this thesis focuses on fundamental biological aspects relevant to the development of improved pest control strategies such as SIT and IIT for Australian tephritid fruit flies, with particular emphasis on the three species *B. tryoni*, *B. neohumeralis* and *B. jarvisi.* The methods used to control pest fruit fly species in Australia are in a state of flux, with environmental and health constraints limiting the use and/or effectiveness of previously utilised chemical control techniques. Bait sprays and lure-and-kill techniques are still viable, but the organophosphate insecticides, dimethoate and fenthion, while providing good control, in particular for post-harvest disinfestation, have been suspended due to their environmental unsustainability and health issues (Dominiak and Ekman 2013).

*Wolbachia* manipulates host reproduction in many insect host species (with CI being an essential requirement for IIT based population suppression). *Wolbachia* can also have physiological effects on host fitness, such as adaptive benefits or detriments for host populations. The search for alternative agents to deliver sterility for IIT (which in conjunction with a male-only strain is an alternative or complementation to SIT) requires an analysis of the incidence and prevalence of *Wolbachia* in Australian species of fruit fly (Chapter 2 and 3) and it could also reveal candidate *Wolbachia* strains that are inherently adapted to Australian tephritid fruit fly species and thus are candidates for introduction of IIT.

Mass-reared fly lines for SIT and IIT are unavoidably affected by domestication of the mass-reared line (Gilchrist *et al.* 2012). However, the general role of microbial symbionts in insects also contributes to host fitness, as well as host immunology, host preference and climatic adaptation. As an application to pest control, the available microbial flora may be manipulated as a method to reduce fitness of fruit flies in the field, or to augment SIT and IIT programmes by improving the rearing methods, fitness and mating performance of mass-reared lines. This study has used for the first time a deep-sequencing approach to build a library of information about the bacterial associations found in natural and laboratory fruit flies, primarily of four *Bactrocera* species: *B. tryoni*, *B. neohumeralis*, *B. jarvisi* and *B. cacuminata* (Chapter 4).

For the development and optimisation of male-only strains of *B. tryoni* and, in future, *B. jarvisi and B. neohumeralis*, the isolation of the essential sex-determination genes, embryonic transcriptome and expression analysis is required. The *Dominant Male Determiner* has not been identified in any insect species thus far and its characterisation is not only important for pest control in *B. tryoni* and other Diptera, but understanding of *M* will greatly augment research into all facets of insect development, gene regulation and evolution. Speculation abounds on what form *M* takes – and includes a small interfering RNA (siRNA), microRNA, a large regulatory RNA or mRNA encoding a protein. Moreover, the target of *M* has yet to be revealed. A detailed examination of the expression patterns of the genes in the *Bactrocera* sex-determination pathway over the early developmental stages using quantitative reverse transcription polymerase chain reaction (qRT-PCR) will shed light on the operation and perhaps interactions of the key genes in this pathway. This will potentially pinpoint at what stage *M* is exerting its influence and highlight the means by which this is achieved. Sequencing of the *Bactrocera* transcriptome at this important time of development will greatly expand our database of embryonic *Bactrocera* sequences and provide access to candidates for *M*.

Therefore, this study will focus on expanding the specific knowledge of sex-determination genes and their expression during early embryonic development in the Australian fruit fly species *B. tryoni* and *B. jarvisi*. This will involve acquiring gene-specific expression data (Chapter 5) and, for *B. jarvisi*, embryonic transcriptome sequence and quantification of gene expression (Chapter 6).

The specific aims were:

1. **To investigate the incidence and prevalence of *Wolbachia* in Australian tephritid fruit fly species (Chapter 2).**
2. **To examine the occurrence of horizontal transmission of *Wolbachia* in Australian tephritid fruit fly communities (Chapter 3).**
3. **To investigate the microbiome of natural and laboratory stocks of**

**tephritid fruit flies (Chapter 4).**

4. **To quantify the expression of sex-determination genes in *B. jarvisi* and *B. tryoni*, using male and female embryos during defined developmental stages (Chapter 5).**

5. **To sequence the poly(A)$^{+}$ fraction of the transcriptome in male and female *B. jarvisi* embryos to increase access to molecular sequence information for this species, and to analyse the differences in gene expression of sex-determination genes and candidate male-specific transcripts (Chapter 6).**

## 1.7 Thesis structure

I have presented this thesis as a series of five experimental papers accepted, submitted or prepared for submission to peer-reviewed journals. The structure of each chapter adheres to the style of the chosen journal. In addition to these five experimental chapters the thesis has been prefaced by an introductory literature review (Chapter 1), and a final Chapter 7 that contextualises the research, discusses key findings and applications, and outlines prospects for future research. Each chapter was co-authored, with the section on molecular isolation of sex-determination genes *tra*, *tra-2* and *Sxl* in Chapter 5 carried out jointly with D.C.A Shearman. Otherwise, I was principally responsible for the concept, experimental design, data collection, analysis and writing of each of the chapters.

**Chapter 2**: Morrow, J. L., Frommer, M., Royer, J., Shearman, D. C. A. and Riegler, M. "Latitudinal cline in the incidence of *Wolbachia* endosymbionts in Australian tephritid fruit fly communities" ISME Journal (in review).

**Chapter 3**: Morrow, J. L., Frommer, M., Shearman, D. C. A. and Riegler, M. "Tropical tephritid fruit fly community with high incidence of shared *Wolbachia* strains as platform for horizontal transmission of endosymbionts" Environmental Microbiology (accepted 23 December 2013).

**Chapter 4:** Morrow, J. L., Frommer, M., Shearman, D.C.A. and Riegler, M. **"**The microbiome of field-caught and laboratory-adapted Australian tephritid fruit fly species" Microbial Ecology (in preparation).

**Chapter 5**: Morrow, J. L., Riegler, M., Frommer, M., and Shearman, D. C. A. "Expression patterns of sex-determination genes in single male and female embryos of two *Bactrocera* fruit fly species during early development" Insect Molecular Biology (in preparation).

**Chapter 6**: Morrow, J. L., Riegler, M., Gilchrist, A. S., Shearman, D. C. A. and Frommer, M. "Transcriptome sequencing of male and female *Bactrocera jarvisi* embryos" BMC Genomics (in preparation)

# Chapter 2

# Latitudinal cline in the incidence of *Wolbachia* endosymbionts in Australian tephritid fruit fly communities

Morrow, J. L., Frommer, M., Royer, J., Shearman, D. C. A. and Riegler, M. “Latitudinal cline in the incidence of *Wolbachia* endosymbionts in Australian tephritid fruit fly communities” ISME Journal (in review).

## 2.1 Abstract

*Wolbachia* bacteria are maternally-inherited endosymbionts that infect a large number of insect species. Besides their well-known role as parasites that highjack host reproductive biology, *Wolbachia* can positively or negatively impact host fitness and, conversely, also respond to environmental conditions of their host species. This has not yet been reflected in analyses of infection frequencies at insect community scales. A limited number of studies have compared incidences of *Wolbachia* in tropical and temperate insect communities, none along an extensive climatic gradient, and none revealed an overall difference. Here, we have performed a field survey of 24 tephritid fruit fly species, collected along a continuing gradient through four climate zones of eastern Australia, to determine their infection status through PCR screening for the *Wolbachia* surface protein gene (*wsp*) and 16S rDNA. *Wolbachia* infections were verified by sequence analysis, while host associations were corroborated through barcoding of mitochondrial *cytochrome oxidase I*. We detected *Wolbachia* in eight out of 24 tephritid fruit fly species. Incidence of *Wolbachia* in fruit fly communities was restricted to northern Australia, including infections in five fruit fly species that only occur in the tropics. *Wolbachia* prevalence within three more widely distributed tephritid species, including the two economically important pests, *Bactrocera tryoni* and *Bactrocera neohumeralis*, was also highest in the tropics. This restriction of *Wolbachia* to the tropical regions of Australia is in contrast to the previously reported global equilibrium of *Wolbachia* infections. It suggests that *Wolbachia* in these fruit fly communities may respond to the different environmental conditions along their latitudinal range, thus creating an opportunity for local adaptation in *Wolbachia*-host interactions. Alternatively, hosts are more frequently exposed to horizontal transmission of *Wolbachia* in tropical regions with higher host species diversity, or *Wolbachia* may be currently invading fruit fly species and populations.

## 2.2 Introduction

*Wolbachia pipientis* (Alphaproteobacteria) is a common endosymbiotic bacterium estimated to infect 40 to 65% of terrestrial arthropod species including insects, arachnids and isopods, as well as filarial nematodes (Hilgenboecker *et al.* 2008, Zug and Hammerstein 2012). *Wolbachia* is mostly maternally inherited, but occasional horizontal transmission into uninfected lineages occurs and contributes to the large number of infected species (Zug *et al.* 2012). Within host species, *Wolbachia* can cause reproductive anomalies including feminisation of genotypic males, mortality of male offspring of infected females, thelytokous parthenogenesis and cytoplasmic incompatibility (CI); all of these result in a reproductive advantage of infected over uninfected females and thus increase the prevalence of *Wolbachia* in host populations (Werren *et al.* 2008). CI is the most commonly reported reproductive phenotype that results in embryonic mortality in crosses between individuals of different infection status (Werren 1997b). Besides its potential in suppression of insect populations (Zabalou *et al.* 2004), *Wolbachia* can also contribute to genetic isolation and speciation of host species (Bordenstein *et al.* 2001, Miller *et al.* 2010).

Fitness effects in *Wolbachia*-infected individuals other than the effect of reproductive manipulation have recently been revealed. Depending on *Wolbachia* and host genotypes, fitness effects can either be beneficial or deleterious to hosts. Examples of positive fitness effects include increased fecundity (Fast *et al.* 2011, Weeks *et al.* 2007), facilitation of host plant use (Kaiser *et al.* 2010), protection from RNA viruses (Hedges *et al.* 2008, Teixeira *et al.* 2008) and support of iron and vitamin B metabolism (Brownlie *et al.* 2009, Hosokawa *et al.* 2010); negative fitness effects include shortened life-span (Min and Benzer 1997), decreased cold tolerance (Maes *et al.* 2012) and increased desiccation susceptibility (McMeniman and O'Neill 2010). Thus, besides its role as reproductive parasite, *Wolbachia* is recognised as an adaptive endosymbiont in an increasing number of host species (Frago *et al.* 2012, Riegler and O’Neill 2007).

Studies highlighting the incidence of *Wolbachia*-infected species in arthropod communities from tropical and temperate regions indicated parity of *Wolbachia* incidence across climatic zones (Ahmed *et al.* 2013, Werren *et al.* 1995a, Werren and Windsor 2000, West *et al.* 1998). In contrast, *Wolbachia* prevalence within

individual host species and populations was shown to vary from very low (Arthofer *et al.* 2009a, Doudoumis *et al.* 2012, Sun *et al.* 2007) to fixation (Riegler and Stauffer 2002). There is evidence for climatic factors involved in shaping some individual *Wolbachia* host associations such as in a leaf beetle species (Keller *et al.* 2004), and this could also be concluded from examination of *Wolbachia* prevalence in *Drosophila melanogaster* populations along the climatic cline of eastern Australia (Hoffmann *et al.* 1994, Hoffmann *et al.* 1998). However, such *Wolbachia* interactions with single host species have not yet been tested for insect communities along a latitudinal gradient.

Here we used the frugivorous community of Australian tephritid fly species collected along a large latitudinal gradient of 3,000km as a model to evaluate climatic signatures in *Wolbachia* infection frequencies. Worldwide, the family Tephritidae encompasses approximately 5,000 species, including key pests of the genera *Anastrepha*, *Bactrocera, Ceratitis*, *Dacus* and *Rhagoletis* (White and Elson-Harris 1992). The largest proportion of Australian tephritids belongs to the genus *Bactrocera* with about 80 endemic species (Drew 1989). The majority originate from and are restricted to the tropical regions (Hancock *et al.* 2000, Royer and Hancock 2012). However, since establishment of horticultural production in Australia in the $19^{th}$ century, several fruit fly species have expanded into more temperate regions, in particular due to planting and invasive expansion of host plants (Meats 1981). The economically most relevant Australian endemic species are *Bactrocera tryoni* and *Bactrocera neohumeralis*, both with an extensive and shared host fruit range from over 40 plant families (Hancock *et al.* 2000).

The main objective of our study was to assess the *Wolbachia* incidence in Australian tephritid fruit fly communities and the *Wolbachia* prevalence within tephritid species across their latitudinal distribution from tropical to temperate regions of eastern Australia. In doing so, we studied *Wolbachia* in insect communities of a continent which, due to its geological history and isolation, has high levels of endemism (Austin *et al.* 2004) and, as such, may also exhibit different *Wolbachia* infection patterns than other continents. We were also interested in characterising the infection status of the economically significant pest *B. tryoni* to see whether it could be targeted by *Wolbachia*-based control strategies such as Incompatible Insect Technique (IIT; Boller *et al.* 1976) in the future.

## 2.3 Materials and methods

*Insect samples*

Both wild and laboratory lines of fruit flies belonging to the genera *Bactrocera*, *Dacus*, *Dirioxa* and *Ceratitis* were analysed. A total of 592 flies from 24 species were field-collected in New South Wales, Queensland, the Northern Territory and the Torres Strait Islands during two sampling periods, from 1996 to 2001, and from 2012 to 2013 (Appendix A: Table A.1). While the majority occur in the equatorial and tropical regions of Australia, ten of the 24 species included in this study also occur in the subtropical and temperate regions (Hancock *et al.* 2000). Most flies were male and collected during summer by trapping with male attractants cue lure (Osborne *et al.* 1997), methyl eugenol for *Bactrocera visenda* and zingerone for *Bactrocera jarvisi* (Royer and Hancock 2012), while both sexes of *B. tryoni*, *Bactrocera cacuminata* and *Dirioxa pornia* were collected directly on or from infested fruit. Fly specimens were identified using identification keys (Drew 1989, White and Elson-Harris 1992). Samples were selected based on availability and to canvas a range of species and populations. At the time of collection, *Bactrocera aquilonis* was recognised as a distinct species that is morphologically similar to *B. tryoni* and found only in the Northern Territory (Morrow *et al.* 2000). This species has since been synonymised with *B. tryoni* (Cameron *et al.* 2010) and here all flies originally classified as *B. aquilonis* have been listed as *B. tryoni.* In addition to field-caught flies, we also screened eight females from each of the following laboratory lines: *B. tryoni*, *B. neohumeralis*, *B. jarvisi*, and *B. cacuminata* kept at UWS, Richmond, New South Wales; two independent *B. neohumeralis* lines from Cairns, Queensland, and *Ceratitis capitata* (Vienna 7/Mix 99) from Perth, Western Australia (Appendix A: Table A.2).

*Screening of tephritid fruit flies for* Wolbachia

PCR-based screening of fruit fly DNA (Appendix A: Table A.3) was undertaken using the *Wolbachia* surface protein (*wsp*) and 16S rRNA loci. Primers for *wsp* were 81F and 691R (Braig *et al.* 1998) or Wsp-F and Wsp-R (Jeyaprakash and Hoy 2000). 16S rDNA was amplified with wspecF and wspecR (Werren and Windsor 2000).

Host mitochondrial *cytochrome oxidase I* (*COI*) fragments were amplified with Dick and Pat (Simon *et al.* 1994) as DNA quality control. Flies were classified as uninfected when repeated attempts with *wsp* and 16S rDNA were negative but *COI* was positive. Amplified *wsp* was subjected to Southern hybridisation with a DIG-labelled *wsp* probe to increase sensitivity and specificity (Arthofer *et al.* 2009b). Individuals were considered *Wolbachia* infected when *wsp* and 16S rDNA primers amplified appropriately sized fragments; *wsp* amplicons hybridised to the *wsp* probe; and *wsp* amplicons produced sequence homologues. When direct sequencing provided evidence of multiple *Wolbachia* strains, cloning and PCR-RFLP sequencing was undertaken. The *COI* locus was sequenced to confirm host species association for *Wolbachia*-positive specimens.

*DNA extraction, PCR, cloning and sequencing*

Genomic DNA was extracted from individual fruit fly abdomens, while the remainder of the specimen was stored in ethanol at -80°C for subsequent independent confirmation of positive results. Prior to DNA extraction, specimens were treated with 4% sodium hypochlorite (Sigma, St Louis, MO) for 5min, then triton-X (0.02%) and then thoroughly rinsed with Milli-Q water to reduce surface contamination. DNA from *D. melanogaster* line *w*1118 (infected with Wolbachia strain wMelPop; Min and Benzer 1997) was used as a positive control. Insect tissue was ground in 1.5mL microcentrifuge tubes with microtube pestles (Scientific Specialities Inc., Lodi, CA) and cell lysis performed overnight followed by extraction according to the GenElute Mammalian Genomic DNA Miniprep kit (Sigma) protocol. Elution of DNA from spin columns was with 100µL nuclease-free water, and 1-2µL was used as template for PCR. Risk of contamination was minimised by routinely replacing stock solutions and dispensing aliquots of stock reagents. Although cross-contamination of flies caught in the same trap or stored in the same tube of ethanol after collection has previously been shown to be unlikely (Duplouy *et al.* 2009), we have further minimised this risk by selecting individuals from different collection sites, surface treatment of samples with sodium hypochlorite prior to DNA extraction plus independent extraction and PCR experiments in different laboratories. All DNA extractions and PCRs were assembled using filter tips to prevent contamination.

Veracity of PCR results was tested by inclusion of no-template controls. All positive amplicons were confirmed by replication and further screening with other primer sets. *COI* and *wsp* amplicons were prepared for direct sequencing by treatment with a combination of 0.5u Exonuclease I (New England Biolabs, Ipswich, MA) and 0.25u Shrimp Alkaline Phosphatase (Promega), with incubation at 37°C for 30min, then 95°C for 5min, prior to sequencing by Macrogen (Seoul, Korea).

For products displaying multiple sequences through direct sequencing, *wsp* was PCR amplified for cloning. Amplicons were either gel-extracted using the Wizard SV Gel and PCR Clean-up System (Promega) and eluted in 25µL nuclease-free water; or used directly in the ligation reaction. Ligation was with 0.5µL pGEM-T Easy vector (Promega), 1X Rapid ligation buffer and 3u T4 DNA ligase (Promega). Transformation of JM109 competent cells (Promega) was according to the manufacturer's protocol. Colonies were smeared into a PCR tube using a sterile pipette tip and subjected to PCR using standard T7 Promoter and SP6 primers with reaction and cycling conditions as described for insect *COI* (Appendix A: Table A.3). Positive clones, recognised by appropriately sized PCR products, were prepared for direct sequencing as described above. A minimum of three clones, but usually eight clones for each transformed ligate were selected for sequencing in both directions, using T7 and SP6 primers.

*DIG Southern hybridisation*

PCR amplicons were also authenticated by Southern hybridisation using DIG DNA Labelling and Detection Kit (Roche Applied Sciences, Indianapolis, IN) based on the higher sensitivity method outlined in Arthofer *et al.* (2009b). The *wsp* probe was generated as described in Table A.3 (Appendix A), using *D. melanogaster w*1118 DNA as template.

*PCR-RFLP*

Single restriction enzyme digestion was performed on *wsp* amplicons to test for multiple infections and to confirm the presence of no more than the two sequence

types revealed via clone sequencing. The sequence differences within two of the *wsp* alleles found as multiple infections enabled *Taq*I (cuts *wsp* allele 661 at position 516) and *Spe*I (cuts *wsp* allele 11 at position 286) to distinguish the alleles. *Taq*I and *Spe*I (Promega) reactions were according to manufacturer's protocols for 3h. The samples were electrophoresed on 1.2% agarose gels. Uncut, *Taq*I cut and *Spe*I uncut bands were independently excised from the gel. Samples were purified using Wizard SV Gel and PCR Clean-up System and sequenced.

*Statistical analyses of* Wolbachia *incidence and prevalence*

*Wolbachia* incidence was defined as the percentage of infected species and *Wolbachia* prevalence as the percentage of infected individuals within a species (Zug and Hammerstein 2012). While we sampled an average of 24 individuals per species, sample size was limited for some species. Thus, we restricted inferences about *Wolbachia* incidence and prevalence to species for which we had at least ten individuals from within the same collection period. This was to reduce the risk of underestimating *Wolbachia* incidences in species where *Wolbachia* occurs at low prevalence and was comparable to other recent studies about infection frequencies (Duron *et al.* 2008, Zug and Hammerstein 2012). Sampling locations were binned into 13 latitudinal groups and, for descriptive purposes, were placed into five climatic regions: equatorial, tropical, subtropical, temperate and grassland. These regions shared similar Köppen climate classifications (Australian Bureau of Meteorology). *Wolbachia* incidence between the northern and southern halves of the gradient was tested through Fisher's exact test on the numbers of species with infected individuals versus species without infected individuals. A linear model (lm) was fitted to the relationship of *Wolbachia* prevalence with individual species distribution, based on the midpoint latitude of each species' geographic range. *Wolbachia* prevalence was tested for species that were represented in at least three latitudinal groups and had evidence of *Wolbachia* infections. In this way six species were included: *B. neohumeralis* (n=132), *B. tryoni* (n=190), *Bactrocera bryoniae* (n=51), *Bactrocera frauenfeldi* (n=34), *Bactrocera strigifinis* (n=37) and *Dacus axanus* (n=10). Infection prevalence across latitude and species was tested using multivariate generalised linear models (manyglm, negative binomial response)

available in the package mvabund specifically designed for multivariate abundance data (Wang *et al.* 2012) in R 2.15 (R Core Team 2012). An analogous analysis was performed on the 2012/2013 sample set of *B. neohumeralis* and *B. tryoni*, to test for a time effect. Fisher's exact test was also applied to test *Wolbachia* prevalence over time within five species that were polymorphic for *Wolbachia* infections and were collected across the two sampling periods.

*Phylogenetic analyses*

DNA sequences were trimmed and edited in Sequencher 4.0 (GeneCodes Corporation) and then analysed in Mega 5.05 (Tamura *et al.* 2011). *COI* and *wsp* genes were independently aligned (MUSCLE algorithm). Pairwise distance matrices were calculated for *COI* using number of differences and p-distance models. Substitution models were selected using Find Best DNA Model (ML), which calculated the lowest Bayesian Information Criterion score (GTR+G for *wsp*; TN93+G+I for *COI*). Bayesian Inference phylogenies were produced by MrBayes 3.2 (Ronquist *et al.* 2012) running $10^7$ generations with a sample frequency of 100. The first 25% of trees were discarded, and a 50% majority rule consensus tree returned.

## 2.4 Results

*Analysis of* Wolbachia *incidence and prevalence*

*Wolbachia* specific primers for *wsp* and 16S rDNA were used to screen 592 field-collected Australian fruit flies representing 24 species of *Bactrocera*, *Dacus* and *Dirioxa.* Overall, individuals of eight species (33%) were positive for both *wsp* and 16S rDNA, and the *Wolbachia* prevalence in these species ranged from 2.1% (4/190) to 100% (5/5; Table 2.1). All individuals from the independently established laboratory lines were negative for *Wolbachia*. Initial screening using *wsp* primers 81F and 691R (Braig *et al.* 1998) appeared to produce false positives for some flies. Thus, primers Wsp-F and Wsp-R (Jeyaprakash and Hoy 2000) were chosen, as their amplicons were more consistent and only occasionally produced spurious bands. Southern hybridisation confirmed specificity as well as improved detection sensitivity. In this way, four additional individuals, one each from four species

**Table 2.1** *Wolbachia* prevalence in six Australian fruit fly species across climate zones. The latitudinal gradient of eastern Australia was divided into groups and represented by a major town or city. Darwin and Alice Springs are included as separate areas. Groups are classified into climate zones according to Köppen classification. Fruit fly species do not occur in grey-shaded regions (Hancock et al. 2000), whereas non-determined (n.d.) regions were not sampled.

| | | Total *Wolbachia* prevalence | | | *Wolbachia* prevalence per climate zone (infected/total) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Equatorial | | Tropical | | | Subtropical | | | | | Temperate | | Grass-lands |
| Fruit fly species | Abbreviation | No. | +ve | % | Torres St | Weipa/ Coen | Darwin | Cairns | Townsville | Mackay | Gladstone | Bundaberg | Brisbane | Lismore | Coffs Harbour | Richmond | Alice Springs |
| | | | | | **10-11°S** | **12-13°S** | **10-12°S** | **16-17°S** | **18-19°S** | **20-21°S** | **22-23°S** | **24-25°S** | **26-27°S** | **28-29°S** | **30-31°S** | **32-34°S** | **23°S** |
| *Bactrocera bryoniae* | *Bb* | 51 | 4 | 7.8 | 0/2 | 1/15 | 0/3 | 3/23 | n.d. | 0/1 | 0/1 | n.d. | 0/5 | 0/1 | n.d. | | |
| *Bactrocera frauenfeldi* | *Bfr* | 34 | 5 | 14.7 | 0/5 | 0/13 | | 5/16 | n.d. | | | | | | | | |
| *Bactrocera neohumeralis* | *Bn* | 132 | 13 | 9.8 | 0/11 | n.d. | | 9/37 | 0/12 | 3/12 | 1/10 | 0/10 | 0/28 | 0/10 | 0/2 | | |
| *Bactrocera strigifinis* | *Bs* | 37 | 5 | 13.5 | 2/10 | 0/12 | | 3/15 | n.d. | | | | | | | | |
| *Bactrocera tryoni* | *Bt* | 190 | 4 | 2.1 | n.d. | 0/13 | 0/6 | 2/40 | 0/10 | 2/12 | 0/12 | 0/10 | 0/37 | n.d. | 0/11 | 0/19 | 0/20 |
| *Dacus axanus* | *Dax* | 10 | 1 | 10 | 0/2 | 1/6 | | 0/2 | n.d. | n.d. | n.d. | | | | | | |
| **Subtotal** | | **454** | **32** | | **2/30** | **2/59** | **0/9** | **22/133** | **0/22** | **5/25** | **1/23** | **0/20** | **0/70** | **0/11** | **0/13** | **0/19** | **0/20** |

(*B. bryoniae*, *B. frauenfeldi*, *B. neohumeralis* and *B. strigifinis*) with very faint *wsp* amplification and undetectable 16S rDNA fragments were confirmed to carry *wsp* DNA, while other individuals in these species were positive for both loci. Southern hybridisation to samples with high titre infections demonstrated that the DIG-labelled *wsp* probe did not bind to primer-dimers or spurious products, but hybridised to *wsp* amplicons that were verified by sequencing.

Spatial analysis revealed that ten of 131 samples from equatorial Queensland (7.6%), 22 of 195 (11.3%) samples from tropical Queensland and the Northern Territory, and 6 of 179 (3.4%) samples from subtropical Queensland were positive for *Wolbachia*. However, none of the 60 samples from temperate New South Wales and none of the 27 flies from central Australia had detectable *Wolbachia* (Appendix A: Table A.1). The most southerly site at which *Wolbachia* was detected was Gladstone (23.88°S), in one *B. neohumeralis* individual. Incidence of *Wolbachia* was higher in the northern half of the gradient (north of Gladstone), with *Wolbachia* incidence in eight of twelve species for which a minimum of ten individuals were tested (Figure 2.1). Seven species from the southern half, with a minimum of ten tested individuals per species, were all negative (Fisher's exact test; $p < 0.05$). A linear model was applied and detected a subtle relationship between the prevalence of *Wolbachia* and the midpoint latitude of species in order to appropriately represent both widespread and more tropically restricted species ($R^2 = 0.46$, $F_{1,9} = 9.33$; $p < 0.05$; Appendix A: Figure A.1). Based on the multivariate analysis of infection prevalence in the six species, *Wolbachia* prevalence was significantly affected by latitude ($p < 0.01$). This was driven by the strong interaction of prevalence by latitude in *B. neohumeralis* and *B. tryoni* detected by univariate analysis with adjusted p-values (p.uni=adjusted) to account for family-wise error across species (Table 2.2A). The cline was further confirmed for the 2012/13 surveys of *B. neohumeralis* and *B. tryoni* ($p < 0.01$), and thus also excluded a change over time in these species (Table 2.2B). Furthermore, temporal effects on *Wolbachia* prevalence within *B. bryoniae, B. frauenfeldi*, *B. neohumeralis*, *B. strigifinis* and *B. tryoni* was tested for equatorial and tropical samples by using a Fisher's exact test ($\alpha = 0.05$) over the sampling years. No significant temporal change within the tested regions was detected, except for *B. tryoni* ($p < 0.05$), but this represented neither an overall increase nor decrease in prevalence over time (Appendix A: Table A.4).

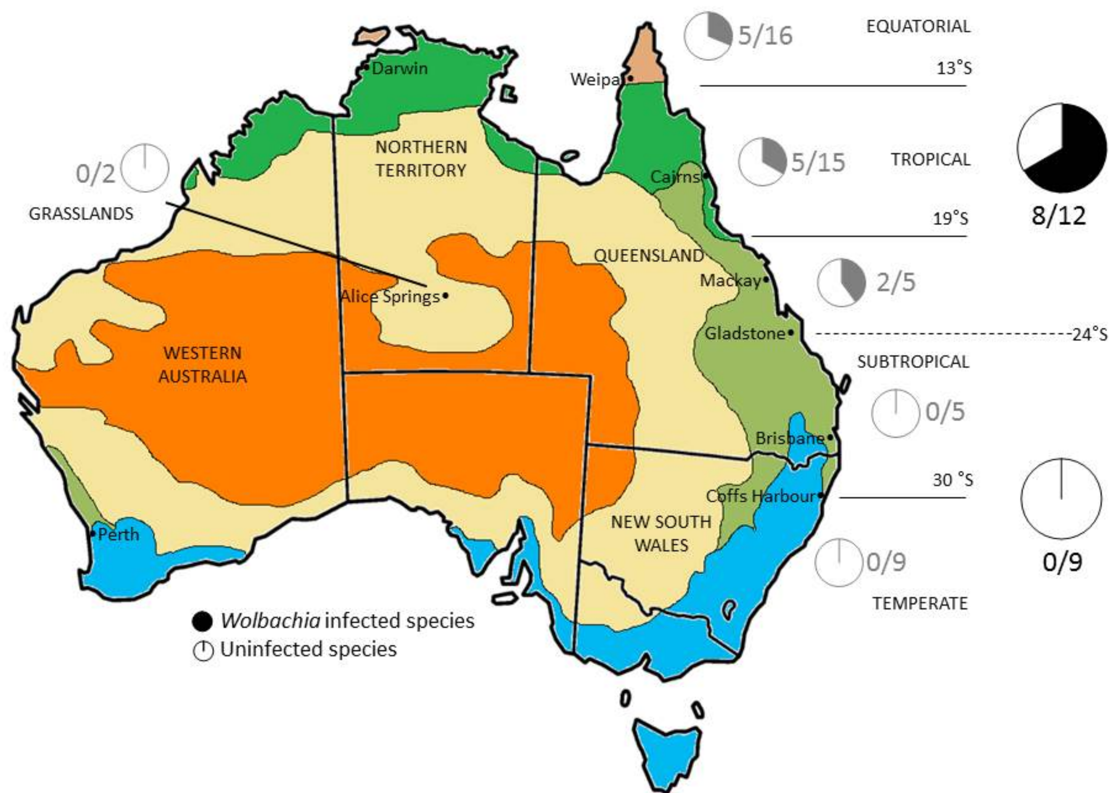



**Figure 2.1** Incidence of *Wolbachia* infection in tephritid species, drawn on the Köppen map of Australia, with climate zones and latitudes as approximate climate divisions along the east coast of Australia. *Wolbachia* incidence (black pie graph; infected species per total number of tested species) is higher in the northern than southern half of the latitudinal gradient (Fisher's exact test; $p < 0.05$). At least ten individuals were tested in the case of absence of *Wolbachia.* Grey pie graphs represent *Wolbachia* incidence in all tested fruit fly species throughout six regions (including species with less than ten individuals per sample).

### *Sequence analysis of* wsp

Thirty-eight of the total 592 flies amplified at the *wsp* locus, and 34 of these also successfully amplified at the 16S rRNA locus. Sequence analysis identified five different *wsp* alleles of A supergroup. Two *wsp* alleles (661 and 11) were found to co-occur in individuals of six species (*B. bryoniae*, *Bactrocera decurtans*, *B. frauenfeldi, B. neohumeralis*, *B. strigifinis* and *D. axanus*); *wsp*-661 occurred

**Table 2.2** Multivariate and univariate analyses of *Wolbachia* abundance (status) at different latitudes of (**A)** six widespread species harbouring *Wolbachia* and (**B)** two widespread species *B. neohumeralis* and *B. tryoni* sampled over 2012/2013 (anova output from mvabund). For species abbreviations, refer to Table 2.1.

**A**

Analysis of Deviance Table — Model: manyglm(formula = mvwolb ~ latitude * status, family = "negative.binomial")

| Multivariate test: | Res.Df | Df.diff | Dev | Pr(>Dev) | |
|---|---|---|---|---|---|
| latitude | 30 | 9 | 128.09 | 0.001 | *** |
| status | 29 | 1 | 191.93 | 0.001 | *** |
| latitude:status | 20 | 9 | 55.75 | 0.001 | *** |

| Univariate Tests: | Bn | | Bt | | Bb | | Bfr | | Bs | | Dax | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Dev | Pr(>Dev) | Dev | Pr(>Dev) | Dev | Pr(>Dev) | Dev | Pr(>Dev) | Dev | Pr(>Dev) | Dev | Pr(>Dev) |
| latitude | 12.15 | 0.375 | 9.56 | 0.384 | 26.33 | 0.032 | 29.64 | 0.012 | 29.72 | 0.012 | 20.69 | 0.086 |
| status | 48.837 | 0.001 | 76.749 | 0.001 | 30.64 | 0.001 | 13.71 | 0.006 | 14.63 | 0.005 | 7.361 | 0.006 |
| latitude:status | 25.734 | 0.001 | 13.378 | 0.021 | 2.466 | 0.657 | 8.52 | 0.067 | 4.559 | 0.434 | 1.095 | 0.657 |

**B**

Analysis of Deviance Table — Model: manyglm(formula = mvwolb ~ latitude * status, family = "negative.binomial")

| Multivariate test: | Res.Df | Df.diff | Dev | Pr(>Dev) | |
|---|---|---|---|---|---|
| latitude | 30 | 9 | 13.9 | 0.607 | |
| status | 29 | 1 | 129.45 | 0.001 | *** |
| latitude:status | 20 | 9 | 29.57 | 0.002 | ** |

| Univariate Tests: | Bn | | Bt | |
|---|---|---|---|---|
| | Dev | Pr(>Dev) | Dev | Pr(>Dev) |
| latitude | 11.03 | 0.456 | 2.87 | 0.88 |
| status | 50.97 | 0.001 | 78.48 | 0.001 |
| latitude:status | 20.18 | 0.003 | 9.39 | 0.011 |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 ; Res. Df = residual degrees of freedom; Dev = deviance

singly in eight individuals (four *B. tryoni,* three *B. neohumeralis* and one *B. frauenfeldi*); and *wsp*-11 in one *B. neohumeralis* individual (Table 2.3, Figure 2.2). PCR-RFLP and sequencing confirmed the absence of any other detectable *wsp* variants in these individuals. Cloning of *wsp* from one *B. frauenfeldi* individual (ID136) revealed two alleles unlike those found in the other *B. frauenfeldi* individuals: one sequence identical to *wsp*-16 of *Drosophila simulans* strain *w*Ri (Baldo *et al.* 2006b); the other (accession no. KC693012) with high similarity to two *wsp* sequences detected in *Bactrocera dorsalis* from China (Sun *et al.* 2007). All five *Bactrocera perkinsi* individuals produced *wsp* and 16S rDNA amplicons. Cloning and sequencing of the *wsp* fragments revealed a novel allele, *wsp*-662, as well as a sequence with a single base insertion, which disrupts the open reading frame by incorporating stop codons.

**Table 2.3** *wsp* allele numbers and GenBank accession numbers from species, confirmed in listed individuals by sequencing and PCR-RFLP.

| Species (double/ single infection) | Individual ID No. | 16SrDNA | wsp† | wsp GenBank accession no. |
|---|---|---|---|---|
| *B. bryoniae* (d) | 157, 536,545 | yes | 11 | KC668327 |
| | | | **661** | KC668326 |
| *B. decurtans* (d) | 85 | yes | 11 | KC668325 |
| | | | **661** | KC668324 |
| *B. frauenfeldi* (d) | 136 | yes | 16 | KC668321 |
| | | | * | KC693012 |
| *B. frauenfeldi* (d) | 485, 492 | yes | 11 | |
| | | | **661** | |
| *B. frauenfeldi* (s) | 490 | yes | **661** | |
| *B. neohumeralis* (d) | 35, 109, 160, 221, 238, 248, 342, 345, 355 | yes | 11 | KC668323 |
| | | | **661** | KC668320 |
| *B. neohumeralis* (s) | 244 | yes | 11 | KC668323 |
| *B. neohumeralis* (s) | 240, 243, 346 | yes | **661** | KC668320 |
| *B. perkinsi* (s) | 74 | yes | **662** | KC668319 |
| *B. strigifinis* (d) | 81, 269, 503, 504 | yes | 11 | KC668329 |
| | | | **661** | KC668328 |
| *B. tryoni* (s) | 275, 276 | yes | **661** | KC668332 |
| *D. axanus* (d) | 88 | yes | 11 | KC668331 |
| | | | **661** | KC668330 |

†*wsp* allele numbers from *Wolbachia wsp* database, new *wsp* alleles in bold; * *wsp* allele number not assigned

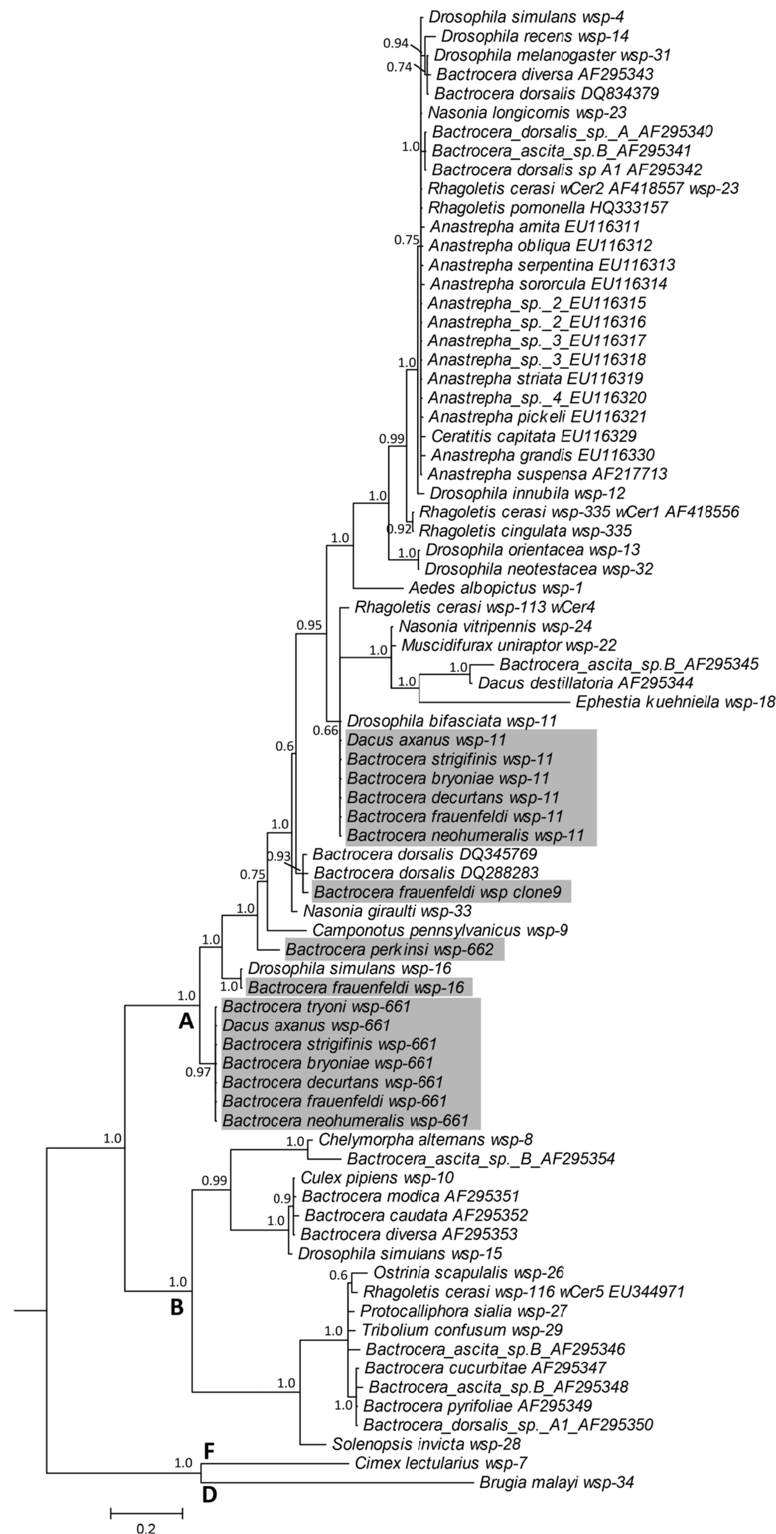



**Figure 2.2** Bayesian inference tree of *wsp* sequences from *Wolbachia* within tephritid flies and other invertebrates. Highlighted sequences are from this study (details in Table 2.3). All other sequences were retrieved from GenBank or the *Wolbachia wsp* database (accession number or *wsp* allele number following host species name). Supergroup classification of *wsp* sequences is marked at the nodes. Posterior probabilities >0.50 are marked at the nodes; tree was rooted by supergroup D and F strains.

*Sequence analysis of 16S rDNA*

The 16S rDNA fragment was consistently amplified from all seven *Bactrocera peninsularis* individuals, whereas the *wsp* locus failed to amplify. Direct sequencing of the 16S rDNA amplicon showed multiple peaks in the sequence chromatogram indicating two distinct sequences: one full length B group 16S rDNA sequence, and one sequence with a 31bp deletion at positions 182-213, indicative of a potential pseudogene. The sequence of the *B. peninsularis* full-length 16S rDNA fragment was deposited in GenBank, accession number KC775793. Failure to detect *wsp* sequence in any of these individuals, suggested that this sequence may not represent a genuine *Wolbachia* infection; consequently this species was classified as uninfected. As a control for the 16S rDNA PCR assays, one *B. neohumeralis* sample, ID248, was chosen for 16S rDNA sequencing to confirm homology to other *Wolbachia* 16S rDNA sequences in GenBank (accession numbers KC775794-KC775795).

*Analysis of mitochondrial DNA*

Most tephritid species produced clear *COI* sequences. *Bactrocera papayae*, *B. jarvisi*, *Bactrocera murrayi* and *B. perkinsi* individuals produced ambiguous sequences, indicative of potential nuclear mitochondrial (numt) DNA in these species. The latter sequences were not included in phylogenetic analysis. Instead, *B. jarvisi* and *B. papayae* sequences were retrieved from GenBank. Bayesian analysis of 81 sequences (52 sequences from this study) over 571bp returned a well-supported consensus tree (Figure 2.3).

Morphologically identified individuals of different species, harbouring identical *wsp* alleles, had different mitochondrial haplotypes. The two sibling species *B. neohumeralis* and *B. tryoni* had individuals with identical *Wolbachia* and mitochondrial sequences but without any linkage between *Wolbachia* infection and haplotypes. *COI* distance measures of *B. neohumeralis* and *B. tryoni* (analysed as one species complex) showed little difference within infected (1.22%) and uninfected individuals (1.36%), and between infected and uninfected individuals (1.31%; Appendix A: Table A.5; Figure 2.3).

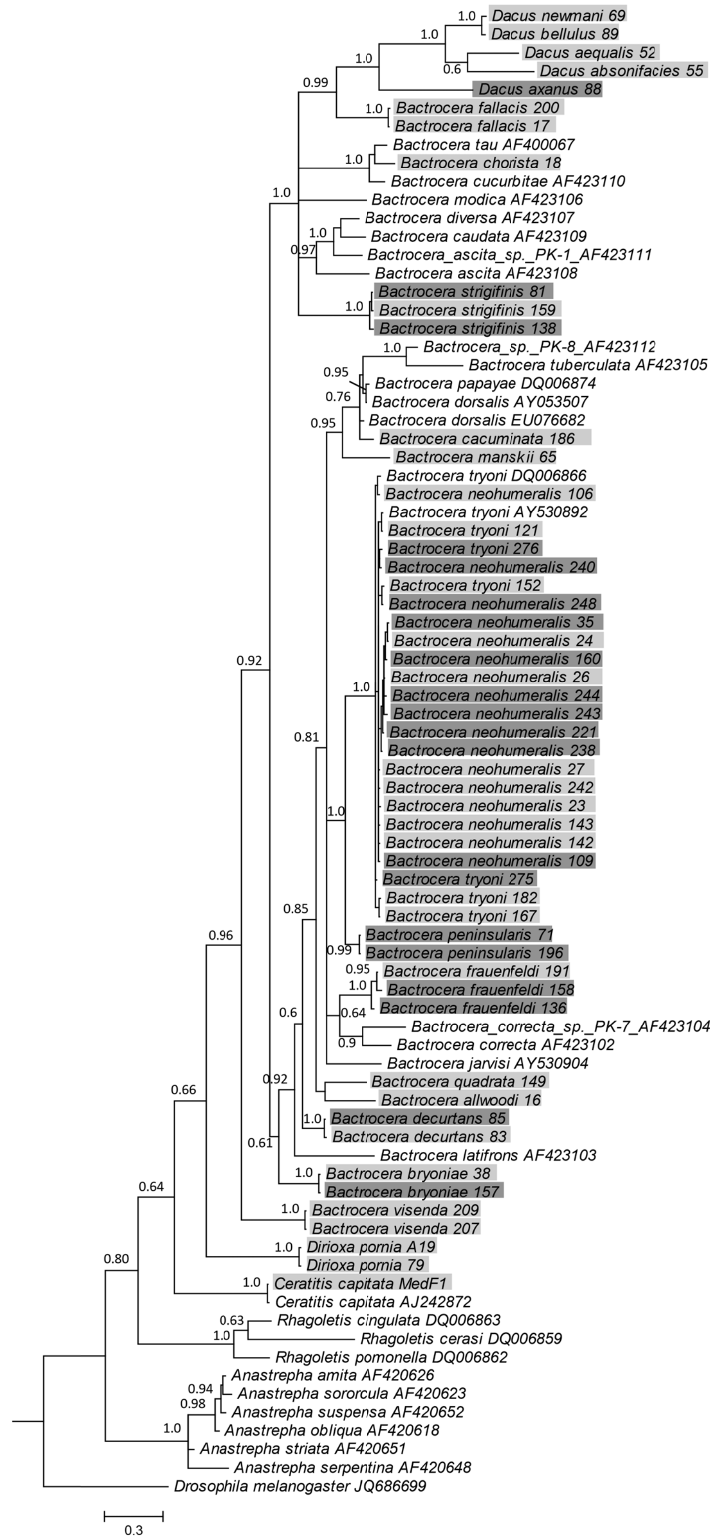



**Figure 2.3** Bayesian inference tree of *COI* sequences from representatives of the five major tephritid fruit fly genera, *Anastrepha, Bactrocera*, *Ceratitis*, *Dacus* and *Rhagoletis* as well as *Dirioxa.* Dark grey shading highlights the individuals from this study harbouring *Wolbachia*, light grey highlights specimens that were negative for *Wolbachia*. Sequences from species without shading were retrieved from GenBank (accession numbers shown). Tree was rooted with *Drosophila melanogaster*, and nodes are labelled with posterior probability values.

## 2.5 Discussion

For the first time, and in contrast to previous studies, we detected a tropically restricted incidence and prevalence of *Wolbachia* in a frugivorous insect community along a large climatic gradient from tropical to temperate regions. We detected *Wolbachia* in eight of 24 Australian tephritid fruit fly species, with an overall low *Wolbachia* prevalence within infected species. *Wolbachia* signals found in two tropical fruit fly species, *B. perkinsi* and *B. peninsularis* could be due to either lateral gene transfer (LGT) from *Wolbachia* into host genomes (Dunning Hotopp *et al.* 2007), pseudogenes within the *Wolbachia* genome, or pseudogenes within some genomes of polyploid *Wolbachia* cells (although there is no published evidence for this alternative). In *B. perkinsi* individuals, some (but not all) *wsp* sequences included stop codons that may be pseudogenes arising from any of the above processes; and LGT to the host genome may be responsible for the amplification of 16S rDNA, but not *wsp* sequences, in *B. peninsularis.* Individuals of six fruit fly species carried two identical *wsp* sequence variants, presumably of two *Wolbachia* strains. Single infections by both of these variants were detected in *B. neohumeralis* individuals and single infections by one variant in all four infected *B. tryoni* individuals and one *B. frauenfeldi* individual.

Measures of *Wolbachia* prevalence are impacted by the sensitivity of the detection method or by *Wolbachia* titre as governed by the age, gender or condition of the sample. High sensitivity Southern hybridisation was used to optimise the detection of low titre infections; despite these measures, low density infections may have escaped detection. Unavoidable variation in age of field collected samples may reflect age-linked titre levels (Arthofer *et al.* 2009b) which can also be specific to the *Wolbachia* and host association, but large sampling efforts can mitigate such variation in detection due to unknown ages of the flies.

### *Latitudinal distribution of* Wolbachia *in Australian tephritids*

We found incidence of *Wolbachia* in 33% of tested tephritid species in Australia, similar to previous estimates of 40% incidence of infections in arthropods, mostly based on a European dataset (Duron *et al.* 2008, Zug and Hammerstein 2012).

However, our results varied between climatic regions. *Wolbachia* incidence was restricted to equatorial and tropical regions north of Gladstone, while *Wolbachia* was absent from individuals caught in most subtropical, temperate and arid inland regions. Specifically, *Wolbachia* was found in species whose range is limited to tropical and equatorial Queensland (*B. decurtans*, *B. frauenfeldi, B. perkinsi, B. strigifinis*); in species with a broader climatic range, *Wolbachia* occurred as far south as Cairns for *B. bryoniae*, Mackay for *B. tryoni* and Gladstone for *B. neohumeralis*, but was absent from individuals in more southerly subtropical and temperate regions.

*Potential drivers for the distribution of* Wolbachia *in Australian tephritids*

Our finding of a latitudinal cline of *Wolbachia* incidence and prevalence in Australian fruit flies may be due to one or a combination of the following scenarios: (1) *Wolbachia* are dependent on the environmental conditions of their hosts and were lost from populations that spread southwards, (2) higher levels of horizontal transmission occur in tropical regions as a result of increased species diversity and exposure to *Wolbachia* of other infected hosts, or (3) *Wolbachia* are currently invading fruit fly species from north to south, e.g. by the means of CI.

An increasing number of studies demonstrate that *Wolbachia* respond to the climatic environment of their host insects. For example, precipitation frequency appeared to correlate with the distribution of single and double infections in leaf beetle individuals in Panama with multiple infections restricted to wetter regions (Keller *et al.* 2004). Multiple infections were more frequent in one tropical habitat when compared with two temperate habitats (Werren and Windsor 2000). Temperature is known to affect *Wolbachia* titres in a number of insect hosts. For example, the virulent *w*MelPop strain appears to over-replicate at elevated temperatures (Reynolds *et al.* 2003). However, high temperatures may also effectively reduce *Wolbachia* densities in hosts through increased *Wolbachia*-bacteriophage activity (Bordenstein and Bordenstein 2011), and thus result in diminished penetrance of CI or male-killing phenotypes (Breeuwer and Werren 1993, Hurst *et al.* 2000). It is possible that temperature fluctuations and extremes, commonly found in southern regions of Australia, may be selecting more strongly against *Wolbachia* than the more constant

temperature conditions in the tropics. Such climatic effects could also be contributing to the distribution of *Wolbachia* in previously reported single species analyses such as *D. melanogaster* (Hoffmann *et al.* 1998) and cat fleas (Tay 2013). Similarly, climate and latitude were found to determine ranges of other microbial symbioses, for example in marine invertebrates (Sanders and Palumbi 2011), terrestrial insects (Dunbar *et al.* 2007, Morag *et al.* 2012, Mueller *et al.* 2011) and humans (Guernier *et al.* 2004). Another hypothesis supposes that the *Wolbachia* in the Australian tephritid species, in particular of the widespread *B. tryoni* and *B. neohumeralis,* is relatively recently acquired through horizontal transmission, and a progressive CI driven sweep may therefore be in its infancy. However this hypothesis is less likely for the following reasons. A CI driven invasion should result in an increase of infection prevalence over time, but this was not detected in these species. Furthermore, *Wolbachia*-induced CI is expected to cause a selective sweep of infected mitochondrial haplotypes (Hurst and Jiggins 2005), yet there does not appear to be support for this in the *B. tryoni* and *B. neohumeralis* species complex (Morrow *et al.* 2000) that both share mitochondrial haplotypes across infected and uninfected individuals. Unfortunately, we were not able to directly assess the CI phenotype of *Wolbachia* in Australian tephritids due to the lack of infected laboratory populations. The CI characterisation of the *Wolbachia* strains will therefore require future field collection efforts, in particular of females, in order to set up infected laboratory colonies for crossing experiments. Testing of field females will also avoid a potential underestimation of *Wolbachia* infection rates due to our male-biased sampling approach that would not detect male-killing *Wolbachia* strains. However, prevalence of male-killing *Wolbachia* may be generally low as found for another fruit fly family, Drosophilidae (Hurst & Jiggins 2000).

*Potential for horizontal transmission of* Wolbachia *between tephritid species*

Our study revealed a high incidence of two shared *wsp* sequence variants in six and seven of 24 Australian tephritid species. This could be due to occasional species hybridisation (and thus a combined mitochondrial and *Wolbachia* introgression), or due to horizontal *Wolbachia* transmission. Our phylogenetic analysis of *COI* in conjunction with the infection status demonstrated that *Wolbachia*-driven

introgression of mitochondrial haplotypes did not occur as infected species had distinct mitochondrial lineages. An exception are the two sibling species *B. neohumeralis* and *B. tryoni* with common mitochondrial haplotypes (Morrow *et al.* 2000). It is thus more likely that horizontal *Wolbachia* transmission into different host lineages has occurred. It will be interesting to further investigate this high potential for horizontal transmission amongst tephritid fruit flies, in particular as species of this frugivorous community are expected to have more species interactions with shared host plants within the tropics (Hancock *et al.* 2000) and thus potentially more *Wolbachia* exposure. The sharing of *Wolbachia* strains appears to be common for *Wolbachia*-tephritid symbioses (Coscrato *et al.* 2009, Riegler and Stauffer 2002, Schuler *et al.* 2009, Schuler *et al.* 2011, Schuler *et al.* 2013). Before a final interpretation can be made about horizontal transmission, the identity of the *Wolbachia* strains of Australian tephritids will need to be fully analysed by Multilocus Sequence Typing (MLST) and Multilocus Variable Number Tandem Repeat Analysis (MLVA) that was established for *Wolbachia* (Baldo *et al.* 2006b, Paraskevopoulos *et al.* 2006, Riegler *et al.* 2012), in combination with strategies that allow allele assignment to strains when multiple infections co-occur in individuals (Arthofer *et al.* 2011).

Our finding of the absence of *Wolbachia* driven introgression of mitochondrial haplotypes in this group of fruit flies also provided a required confirmation of previous phylogenetic analyses that were mostly based on mitochondrial gene sequences (Krosch *et al.* 2012). We also found evidence that the genus *Bactrocera* is not monophyletic, as species of the subgenus *Zeugodacus* such as *B. strigifinis* clustered with Australian *Dacus*. Blacket *et al*. (2012) have previously detected numtDNA amplicons by using primers for a different section of *COI*; our different primer set did not yield numtDNA for *B. tryoni* and *B. neohumeralis* but revealed potential numtDNA sequences in other species such as *B. jarvisi*, *B. papayae, B. murrayi* and *B. perkinsi*.

*Application of* Wolbachia *for the control of Australian tephritids*

One of the aims of this study was to determine if *B. tryoni* was infected with *Wolbachia* in the field, ascertaining the feasibility of *Wolbachia*-induced IIT. Only a

small proportion of *B. tryoni* (2.1%) flies were infected, and none from the southern range where IIT could be employed as a stand-alone technique or in conjunction with the sterile insect technique (SIT; Zabalou *et al.* 2009), which has been established for *B. tryoni* (Andrewartha *et al.* 1967). Artificial transfer of CI causing strains by microinjection (Riegler *et al.* 2004) has been successfully achieved in the fruit flies *C. capitata* (Zabalou *et al.* 2004) and *Bactrocera oleae* (Apostolaki *et al.* 2011), and this method could be used to introduce a novel *Wolbachia* infection into *B. tryoni* with the aim to cause CI between infected laboratory-reared males and field females that are either uninfected or infected with an incompatible strain.

## 2.6 Conclusions

The restriction of *Wolbachia* in Australian tephritid fruit flies to more tropical regions was an unexpected finding. Previous surveys undertaken on phylogenetically diverse insect communities from one (e.g. Duron *et al.* 2008, Zug and Hammerstein 2012) or several continents but without a gradient (e.g. Werren and Windsor 2000) either assumed or suggested a global equilibrium of *Wolbachia* infection frequencies in insect communities. A recent survey of fig wasp species revealed a higher but equal incidence of *Wolbachia* across four different continents (Ahmed *et al.* 2013), however in absence of spatial or climatic analysis. It is possible that our findings of a *Wolbachia* cline are unique to the group of tephritid fruit flies in Australia, but our study also indicates that *Wolbachia* infection frequencies cannot be generalised across insect communities and spatial scales. Furthermore, climatic gradients that impact *Wolbachia* incidence and prevalence may create feedbacks such as the restriction of *Wolbachia* phenotypes to certain environments and thus create the opportunity for local adaptation in *Wolbachia*-host interactions. Australian tephritid fruit flies may be useful models to investigate this further.

# Chapter 3

# Tropical tephritid fruit fly community with high incidence of shared *Wolbachia* strains as platform for horizontal transmission of endosymbionts

Morrow, J. L., Frommer, M., Shearman, D. C. A. and Riegler, M. "Tropical tephritid fruit fly community with high incidence of shared *Wolbachia* strains as platform for horizontal transmission of endosymbionts" Environmental Microbiology (accepted December 2013).

## 3.1 Summary

*Wolbachia* are endosymbiotic bacteria that infect 40 to 65% of arthropod species. They are primarily maternally inherited with occasional horizontal transmission for which limited direct ecological evidence exists. Previously, we detected *Wolbachia* in eight out of 24 species of Australian tephritid fruit flies. Here, we have used Multilocus Sequence Typing (MLST) to further characterise these *Wolbachia* strains, plus a novel quantitative PCR method for allele assignment in multiple infections. Based on five MLST loci and the *Wolbachia* surface protein gene (*wsp*), five *Bactrocera* and one *Dacus* species harboured two identical strains as double infections; furthermore, *Bactrocera neohumeralis* harboured both of these as single or double infections, and sibling species *Bactrocera tryoni* harboured one. Two *Bactrocera* species contained *Wolbachia* pseudogenes, potentially within the fruit fly genomes. A fruit fly parasitoid, *Fopius arisanus* (Braconidae) shared identical alleles with two *Wolbachia* strains detected in one *Bactrocera frauenfeldi* individual. We report an unprecedented high incidence of four shared *Wolbachia* strains in eight host species from two trophic levels. This suggests frequent exposure to *Wolbachia* in this tropical tephritid community that shares host plant and parasitoid species, and also includes species that can hybridise. Such insect communities may act as horizontal transmission platforms that contribute to the ubiquity of the otherwise maternally-inherited *Wolbachia*.

### 3.2 Introduction

*Wolbachia* (Alphaproteobacteria) are common endosymbionts of invertebrates (Werren *et al.* 2008). Although typically transmitted from mother to offspring through the egg cytoplasm, there are clear indications that horizontal transmission must occur between individuals, both within and across taxonomic groups. Strict maternal inheritance of *Wolbachia* should result in concordant host mitochondrial and symbiont phylogenies, as seen in filarial nematodes that are dependent on *Wolbachia* for their development and reproduction (Casiraghi *et al.* 2001). However, *Wolbachia* and host phylogenies are rarely in agreement for arthropods (Werren *et al.* 2008) where *Wolbachia* bacteria are, with an estimated 40 to 65% of infected species, extraordinarily pervasive (Hilgenboecker *et al.* 2008, Zug and Hammerstein 2012). In many arthropod host species, *Wolbachia* bacteria induce reproductive modifications that bestow advantages on infected over uninfected females, and thus often increase their prevalence in host populations. Cytoplasmic incompatibility (CI) is the most commonly reported reproductive manipulation, and can result in non-viable embryos in matings between infected males and uninfected females, or between individuals infected by incompatible *Wolbachia* strains, while individuals infected with the same strains are compatible (Werren 1997b). *Wolbachia* bacteria also induce aberrations of host reproduction such as feminisation of genotypic males, killing of male offspring of infected females and thelytokous parthenogenesis (Werren *et al.* 2008).

Relatively few examples of phylogenetic concordance of *Wolbachia* bacteria and arthropods are reported (Werren *et al*. 1995) while phylogenetic incongruence is frequently shown by the detection of very similar or identical *Wolbachia* strains among different *Drosophila* species (Baldo *et al.* 2006b, Haine *et al.* 2005, Miller and Riegler 2006), as well as in more taxonomically diverse insects that share host plants (Sintupachee *et al.* 2006), ecological communities (Kittayapong *et al.* 2003), parasitoid-host interactions (Rozhok *et al.* 2011, Vavre *et al.* 1999, Werren *et al.* 1995b) and predator-prey interactions (Hoy and Jeyaprakash 2005, Kittayapong *et al.* 2003). Experimentally, horizontal transfer of *Wolbachia* between hosts by microinjection of donor cytoplasm into uninfected embryos has been successful within and between species (Boyle *et al.* 1993), genera (Zabalou *et al.* 2004) and families (Braig *et al.* 1994, McMeniman *et al.* 2009, Riegler *et al.* 2004). Horizontal

transfer of *Wolbachia* between individuals was also successful when hemolymph was transferred between terrestrial isopod species (Rigaud and Juchault 1995), and occurred between host and parasitoid (Heath *et al.* 1999), and within and between parasitoid species that developed within the same host (Huigens *et al.* 2000, Huigens *et al.* 2004).

*Wolbachia* surveys in the family of tephritid fruit flies, including the genera *Anastrepha, Bactrocera, Ceratitis, Dacus* and *Rhagoletis*, revealed strains belonging to supergroups A or B (Coscrato *et al.* 2009, Jamnongluk *et al.* 2002, Riegler and Stauffer 2002, Rocha *et al.* 2005), which also contain the majority of *Wolbachia* strains of insects (Werren *et al.* 2008). Tephritids appear to be exceptional in their potential for co-infections by multiple *Wolbachia* strains, with reported incidences of five strains in individuals of *Bactrocera ascita* (Jamnongluk *et al.* 2002) and *Rhagoletis cerasi* (Arthofer *et al.* 2009b). According to sequence analyses of the *wsp* gene, some tephritid species share identical (Schuler *et al.* 2009) or similar alleles (Coscrato *et al.* 2009, Schuler *et al.* 2011). However, characterisation of just a single *Wolbachia* locus such as *wsp* lacks strain resolution due to the high recombination rate documented for *Wolbachia* genomes (Baldo *et al.* 2006a, Klasson *et al.* 2009) and may lead to an overestimation of horizontal transmission rates. Therefore, the identity of *Wolbachia* strains needs to be analysed further with approaches such as Multilocus Sequence Typing (MLST; Baldo *et al.* 2006b, Paraskevopoulos *et al.* 2006) or Multiple Locus Variable Number Tandem Repeat Analysis (MVLA; Riegler *et al.* 2012) in combination with strategies that allow allele assignment to strains when multiple infections co-occur in individuals (Arthofer *et al.* 2011). Only a few studies have shown, by MLST characterisation, convincing evidence for identical (Baldo *et al.* 2008, Salunke *et al.* 2012, Stahlhut *et al.* 2010) or multiple identical *Wolbachia* MLST sequence types (ST) in different host species (Schuler *et al.* 2013) and, thus, recent horizontal transmission.

Australia has over 80 endemic species of *Bactrocera* and *Dacus* (Drew 1989), including economically relevant *Bactrocera tryoni* and *Bactrocera neohumeralis*, both with an extensive and shared host fruit range from over 40 plant families (Hancock *et al.* 2000). *Bactrocera tryoni* in particular, with its wide climatic adaptation potential and geographic distribution (Meats 1981, Yonow and Sutherst 1998), coupled with an extensive host range, is a highly invasive species. In a recent

study (Chapter 2), we detected *Wolbachia* sequences in eight out of 24 (33%) species of the tephritid subfamily Dacinae in tropical regions of Australia. *Wolbachia* occurred at low prevalence in *B. tryoni, B. neohumeralis*, *Bactrocera strigifinis*, *Bactrocera bryoniae*, *Bactrocera decurtans*, *Bactrocera frauenfeldi* and *Dacus axanus*; and at high prevalence in *Bactrocera perkinsi*. Co-infections of *Wolbachia* bacteria within fruit fly individuals were found in six of these eight species, and seven species shared at least one identical *wsp* sequence variant.

We have now fully characterised the *Wolbachia* strains from the eight infected tephritid fruit fly species and, for the first time, from two of their parasitoid species, using the MLST approach and a novel allele assignment technique by quantitative allele-specific PCR. Our aim was to determine the extent to which *Wolbachia* strains are shared between taxa within this community. Based on the previously detected identity of the *wsp* gene in seven Australian dacine species, the flies' phylogenetic relatedness, potential for hybridisation and shared ecological niche, including the sharing of food sources and parasitoids, we predicted that some fruit fly species would carry highly similar or identical *Wolbachia* strains. We also tested two common parasitoid species of tephritid fruit flies to determine whether they carry similar or identical *Wolbachia* strains. The results have shown an unprecedented high incidence of shared *Wolbachia* strains within the Australian community of tropical tephritid fruit flies.

## 3.3 Results

*PCR screening of fruit flies*

One hundred and four tephritid fruit fly total genomic DNA extracts of male flies, previously screened for the presence of *Wolbachia*-specific loci *wsp* and 16S rDNA, were subjected to further screening with standard *Wolbachia*-specific MLST primers, as well as their nested primer sets or B group-specific primer sets, to corroborate the infection status (Appendix B: Table B.1). Primers designed for quantitative PCR (Appendix B: Table B.2), which amplified shorter fragments of some of these loci, were also utilised for screening individuals from species that were known to harbour *Wolbachia* strains. These primers would potentially detect *Wolbachia* DNA that,

through degradation, had failed to amplify using primers producing a larger amplicon. Using this technique, we confirmed the infection status of four flies with a double infection that had amplified only weakly at the *wsp* locus, but not at the 16S rRNA locus or at MLST loci with the standard primer sets.

*MLST characterisation by sequence analysis*

MLST profiling was carried out for individuals from eight tephritid fruit fly species for which infection status was confirmed, plus a ninth species, *Bactrocera peninsularis*, that had not amplified *wsp* but contained one intact variant of *Wolbachia* 16S rDNA (GenBank accession number KC775793) and one that appeared to be a 16S rRNA pseudogene. Sequences from the nine fruit fly species, derived from both cloned PCR products and direct amplicon sequencing, were submitted to the *Wolbachia* MLST database (Table 3.1). One *B. neohumeralis* individual (ID244) carried a single novel *Wolbachia* sequence type ST-289; three *B. neohumeralis*, one *B. frauenfeldi* and four *B. tryoni* specimens carried novel ST-285 alone. In addition, nine *B. neohumeralis*, as well as *B. bryoniae*, *B. decurtans, B. frauenfeldi*, *B. strigifinis* and *D. axanus* individuals, possessed both ST-285 and ST-289 as double infections. Another *B. frauenfeldi* fly (ID136) harboured two *Wolbachia* strains that did not share any alleles with either ST-285 or ST-289, but with ST-17 and ST-370 from the MLST database (Table 3.1). The veracity of these results was confirmed by extracting DNA from the remaining thorax of individuals from each species, using new reagents for extraction and PCR, and carrying out the procedures in a different laboratory.

All five *B. perkinsi* individuals successfully amplified at all loci except *ftsZ*. Cloning and sequencing of the four remaining MLST markers of *B. perkinsi* was required because direct amplicon sequencing produced ambiguous sequence chromatograms for all loci except for *gatB*. Sequences of cloned fragments revealed two types of amplicons: complete, coding fragments of open reading frames (ORFs) of the MLST markers, and incomplete fragments with large deletions and frameshifts. For *coxA*, two fragment types were cloned, one of 461bp with a 25bp deletion, and one 486bp; in *hcpA* two deletion fragments were found of 449bp and 476bp (with 65bp and 38bp deletions) as well as the standard 514bp fragment; and *fbpA* amplification yielded a

**Table 3.1** MLST and *wsp* alleles from nine tephritid species and one parasitoid *Fopius arisanus.*

| Species (Abbreviation) | *Wolbachia* prevalence infected/total (%) | Individual ID No. | alleles | | | | | | ST | strain |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | *wsp* | *gatB* | *coxA* | *hcpA* | *ftsZ* | *fbpA* | | |
| *B. bryoniae (Bb)* | 4/51 (7.85%) | 157, 146, 536, 545 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Bbry_A_1 |
| | | | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Bbry_A_2 |
| *B. decurtans (Bd)* | 1/6 (16.7%) | 85 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Bdec_A_1 |
| | | | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Bdec_A_2 |
| *B. frauenfeldi (Bf)* | 5/34 (14.7%) | 485, 492 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Bfra_A_1 |
| | | | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Bfra_A_2 |
| | | 490 | **661** | | 84 | | **169** | 62 | **285** | |
| | | 136 | 16,† | 22,87 | 23,111 | 24,103 | 3,70 | 23,186 | | |
| *B. neohumeralis (Bn)* | 13/132 (9.8%) | 35, 109, 221, 238, 248, 342, 345, 355 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Bneo_A_1 |
| | | | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Bneo_A_2 |
| | | 244 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Bneo_A_1 |
| | | 240, 243, 346 | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Bneo_A_2 |
| *B. peninsularis (Bpen)* | 0/7 (0%) | 71, 72, 73, 196 | | | | | | **196** | | |
| *B. perkinsi (Bper)* | 5/5 (100%) | 74, 75, 76, 261, 263 | **662,*** | **206** | **193,*** | §,* | | §,* | | |
| *B. strigifinis (Bs)* | 5/37 (13.5%) | 81, 269, 503, 504 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Bstr_A_1 |
| | | | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Bstr_A_2 |
| *B. tryoni (Bt)* | 4/190 (4.1%) | 275, 276, 439, 443 | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Btry_A |
| *D. axanus (Dax)* | 1/10 (10%) | 88 | 11 | **204** | 15 | **218** | 6 | 17 | **289** | Daxa_A_1 |
| | | | **661** | **205** | 84 | 68 | **169** | 62 | **285** | Daxa_A_2 |
| *Fopius arisanus (Far)* | 9/9 (100%) | P3 | 16,# | 22,87 | 23,111 | 24,103 | 3,70 | 23,186 | | |
| | | P11, P12, P13, P14, P15, P16, P17^ | 16,# | | | | | | | |

Prevalence data and *wsp* sequences, except for *Fopius arisanus*, are from Chapter 2; † clone sequences obtained for two strains (one sequence is identical to *wsp* 16 of *w*Ri, the second *wsp* sequence was not given an allele number); # clone sequences not obtained, but direct sequencing demonstrates presence of a second allele; ^ *wsp* sequences were confirmed for seven *F. arisanus*; § clone sequences obtained but not given an allele number; * sequences with indels causing frameshifts, possibly representing pseudogenes; new alleles and sequence types (ST) are in bold.

488bp fragment with a frameshift and the expected 508bp coding ORF fragment. Sequencing of the *wsp* amplicon had similarly revealed two sequences: a coding ORF and a fragment with a single base insertion. The *hcpA*, *fbpA* and *wsp* fragments with indels incorporated stop codons due to the frameshift, and may represent pseudogenes integrated into host chromosomes, however this was not further analysed.

For *B. peninsularis*, only 16S rDNA and B group-specific *fbp*A primers successfully amplified for all six individuals, while no other MLST markers or *wsp* were successful. A single novel *Wolbachia fbpA* ORF fragment was retrieved from this species (Table 3.1).

Parasitoids of tephritid fruit flies were added to this analysis due to their potential role as vectors for horizontal transmission of *Wolbachia*. Parasitoid wasps were collected in an orchard in Richmond, New South Wales (a region with only uninfected tephritid fruit fly species), and from wild tobacco (*Solanum mauritianum*) plants in Atherton, Queensland (Table 3.2). Furthermore, fruit fly larvae were collected from infested wild tobacco fruits at this latter site in Queensland, and determined to be *Bactrocera cacuminata*, by *COI* barcoding (Table 3.3). The parasitoids were first grouped into two morphospecies, and then characterised by DNA sequencing. Fifteen parasitoids were grouped into one morphotype with a BLAST search match to mitochondrial *COI* sequence of *Diachasmimorpha tryoni* (Spinner *et al.* 2011) and the remaining nine individuals of the second morphotype with a BLAST search match to *Fopius arisanus* (Quimio and Walter 2001); both are parasitoids (Hymenoptera, Braconidae) of tephritid fruit flies including *B. tryoni* and *B. neohumeralis* (Carmichael *et al.* 2005). None of the 15 *Diachasmimorpha tryoni* wasps (both females and males) but all *F. arisanus* wasps (both females and males) were positive for *Wolbachia*. Complete MLST analysis of one *F. arisanus* individual revealed two alleles at each locus (Table 3.1). One *B. frauenfeldi* individual (ID 136) was found to harbour this same combination of alleles, with one MLST allele at each locus displaying identity with alleles of *w*Ri (ST-17) from *Drosophila simulans* (Baldo *et al.* 2006b) and one identical to *w*Ajap (ST-370) of the drosophilid parasitoid *Asobara japonica* (Kraaijeveld *et al.* 2011). Furthermore, one *B. cacuminata* larva possessed the same two *wsp* alleles as *F. arisanus* and *B. frauenfeldi* (ID136) but was not tested for other MLST alleles; no other

**Table 3.2** Parasitoid wasps used in this study were collected from Richmond, N.S.W. (33.60°S, 150.75°E) or Atherton, Qld (17.26°S, 145.49°E). Specimens were divided morphologically into two genera; two regions of *COI* sequence determined species groupings. Individuals were screened with *wsp*, 16S rDNA primers and sequenced with MLST primers.

| | Collection | | | | | | MLST | | COI GenBank acc. No. | |
|---|---|---|---|---|---|---|---|---|---|---|
| Species | Year | Site | Gender M/F | ID No. | *wsp* | 16S | Standard (F1-R1) | Nested (F3-R3/F1-R1) | Dick-Pat | LCO1490-HCO2198 |
| *Diachasmimorpha tryoni* | 2012 | Richmond | F | P1 | neg | neg | neg | neg | KC581416 | KC857545 |
| *Diachasmimorpha tryoni* | 2012 | Richmond | F | P2 | neg | neg | neg | neg | KC581415 | KC857546 |
| *Diachasmimorpha tryoni* | 2010 | Richmond | F | P5 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2010 | Richmond | F | P6 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2010 | Richmond | F | P7 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2010 | Richmond | F | P8 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2010 | Richmond | F | P9 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | M | B5 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | F | B6 | neg | neg | neg | neg | KC581413 | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | F | B7 | neg | neg | neg | neg | KC581413 | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | F | B8 | neg | neg | neg | neg | KC581413 | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | M | B9 | neg | neg | neg | neg | KC581414 | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | M | B10 | neg | neg | neg | neg | | KC857547 |
| *Diachasmimorpha tryoni* | 2012 | Richmond | M | B11 | neg | neg | neg | neg | | |
| *Diachasmimorpha tryoni* | 2012 | Richmond | M | B12 | neg | neg | neg | neg | | |
| *Fopius arisanus* | 2012 | Richmond | F | P3 | pos | pos | pos | | KC581417 | KC857548 |
| *Fopius arisanus* | 2010 | Richmond | F | P10 | pos | pos | | | | |
| *Fopius arisanus* | 2013 | Atherton | M | P11 | pos | pos | pos | | n | y |
| *Fopius arisanus* | 2013 | Atherton | M | P12 | pos | pos | pos | | n | y |
| *Fopius arisanus* | 2013 | Atherton | M | P13 | pos | pos | pos | | n | y |
| *Fopius arisanus* | 2013 | Atherton | F | P14 | pos | pos | pos | | n | y |
| *Fopius arisanus* | 2013 | Atherton | F | P15 | pos | pos | pos | | n | y |
| *Fopius arisanus* | 2013 | Atherton | F | P16 | pos | pos | pos | | n | y |
| *Fopius arisanus* | 2013 | Atherton | F | P17 | pos | pos | pos | | n | y |

**Table 3.3** Larvae collected from wild tobacco plants in Atherton, Qld (17.26°S, 145.49°E), identified as *B. cacuminata* by *COI* sequence homology to one adult *B. cacuminata* from Richmond, NSW (33.60°S, 150.75°E), and screened for *Wolbachia wsp* and 16S rDNA.

| Species | Collector and year | Location | ID No. | *wsp* | 16S rDNA | COI sequence homology (%) |
|---|---|---|---|---|---|---|
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath1 | no | no | 99.8 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath2 | no | no | 99.6 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath3 | no | no | 99.6 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath4 | no | no | 99.6 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath5 | no | no | 99.8 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath6 | yes | yes | 99.8 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath7 | no | no | 99.8 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath8 | no | no | 99.8 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath10 | no | no | 99.8 |
| *Bactrocera cacuminata* | MR2013 | Atherton | Ath11 | no | no | 100 |
| *Bactrocera cacuminata* | MR2013 | Richmond | 186 | no | no | KC581374 |

*B. cacuminata* adult or larva (Table 3.3; Appendix A: Table A.1) was found to harbour any *Wolbachia*. The detection of these two *Wolbachia* strains in one individual *B. cacuminata* larva may have resulted from carry-over of *Wolbachia* or its DNA following parasitism or injury by a parasitoid, or it may represent a heritable infection. The presence of mixed host and parasitoid DNA can be detected by PCR (Arthofer *et al.* 2009a), however, direct sequencing of the *COI* gene showed the expected *B. cacuminata* sequence but no detectable background sequence to indicate the presence of parasitoid DNA.

We tested the probability of deriving ST-17 and ST-370, detected in the parasitoid *F. arisanus*, in a *B. frauenfeldi* adult and potentially in a *B. cacuminata* larva, by binomial sampling of sequence types from the MLST database (with 645 isolates of 362 sequence types). It is highly unlikely that the same ST was independently isolated from three of the 26 tested host species in our study (binomial test for ST-17; $p < 0.0001$; binomial test for ST-370; $p < 0.0001$) supporting the common source of *Wolbachia* in both parasitoids and hosts. This was also confirmed for individual MLST allele combinations. While allele *ftsZ*-3 of ST-17 was shared with 64 other *Wolbachia* isolates, all other MLST alleles of ST-17 and ST-370 were rare so that it was unlikely that they were independently detected in three different host species (Appendix B: Table B.3).

*MLST allele assignment by quantitative PCR*

*Bactrocera tryoni, B. neohumeralis* and *B. frauenfeldi* included individuals harbouring single strains of *Wolbachia*, allowing the assignment of alleles to particular strains in doubly-infected *B. neohumeralis* and *B. frauenfeldi* individuals by the allele intersection method (Arthofer *et al.* 2011). This led to the complete MLST characterisation of two *Wolbachia* strains in these host species with sequence type designations ST-285 and ST-289, described above. Corroboration of the sequence types in *B. bryoniae*, *B. decurtans*, *B. strigifinis* and *D. axanus* was performed using a novel quantitative PCR approach, as infections only occurred as double infections. This method measured the relative density of *Wolbachia* alleles to each other compared to a host gene, and linked the alleles of a strain by their titres. This approach also allowed relative quantification of *Wolbachia* densities across their five doubly-infected host species plus the singly-infected individuals of *B. neohumeralis* and *B. tryoni*. The quantification cycle (Cq) values were calculated for eight *B. neohumeralis* individuals, two *B. tryoni*, two *B. strigifinis*, and one individual of each *B. decurtans*, *B. bryoniae*, and *D. axanus*, and then normalised by the $2^{-\Delta Cq}$ equation against *scarlet*, a single copy nuclear gene of the fly genome (Appendix B: Table B.4).

ANOVA of the Cq values demonstrated different densities of *Wolbachia* strains found in individuals within species and across species based on *wsp*-11 ($F_{10,25}$ = 1688, $p < 2.2e-16$) and *wsp*-661 ($F_{13,37} = 152.7$, $p < 2.2e-16$), confirmed by Tukey contrasts (Appendix B: Table B.4; Figure 3.1). The range of densities in the eight examined *B. neohumeralis* individuals for the strain carrying *wsp*-11 was from 0.0008 to 6.25 (7,688-fold difference) and for *wsp*-661 from 0.0013 to 2.21 (1,657-fold difference). Furthermore, based on the *wsp* locus, the ratio of *wsp*-11 to *wsp*-661 ranged from 4.29 to 0.61, indicating variation in the stoichiometry of the two *Wolbachia* strains across species.

The qPCR analyses were primarily directed toward linking the STs found in single infections of *B. tryoni* and *B. neohumeralis* with the double infections harboured by *B. bryoniae*, *B. decurtans*, *B. neohumeralis*, *B. strigifinis* and *D. axanus*, to determine if recombination had produced different strains in these hosts. Implementation of

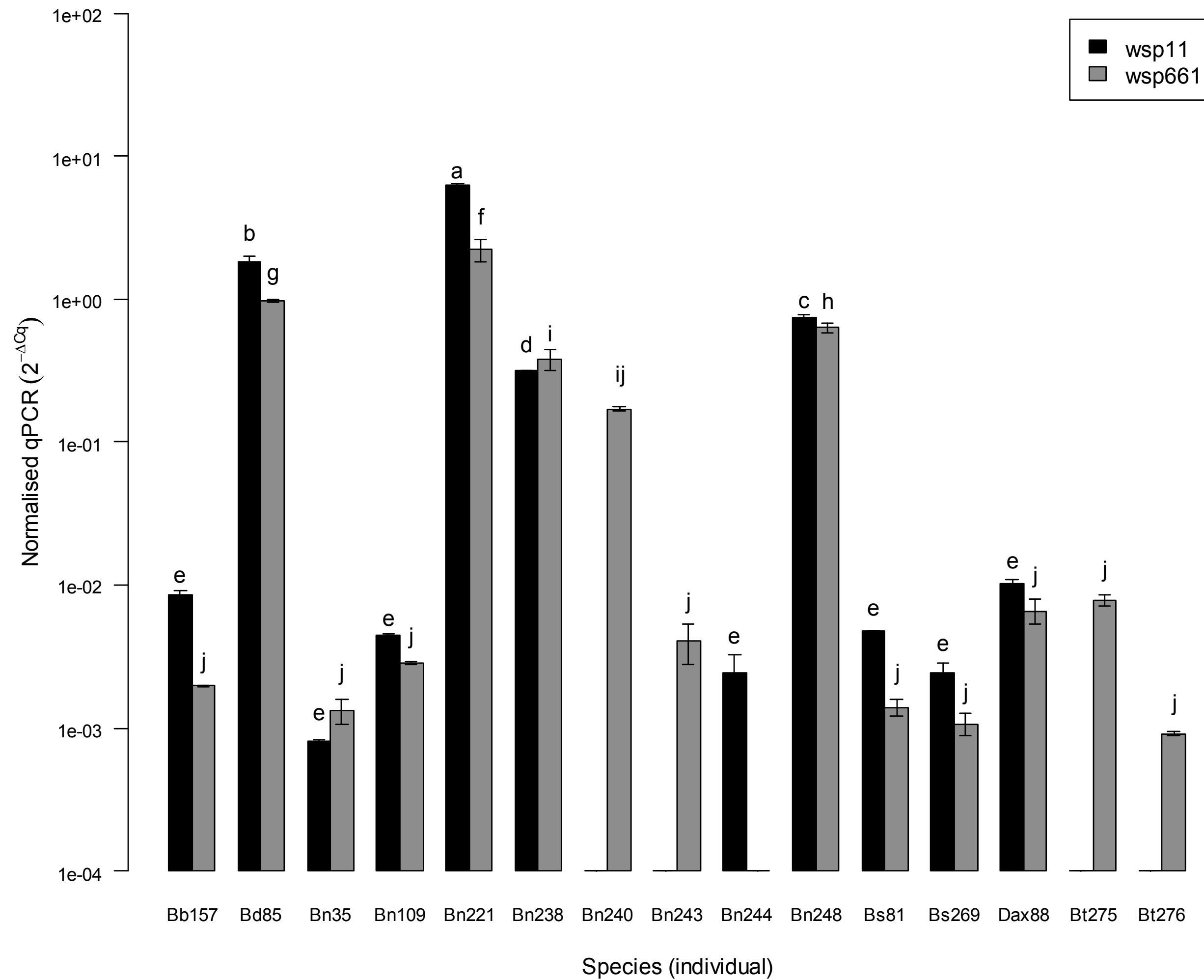



**Figure 3.1** Comparison of *wsp* alleles' overall titres, and relative difference of strain titre within and across species based on *wsp* alleles using quantitative PCR. Quantification cycle values (Cq) were normalised (see Appendix B: Table B.4) and graphed on a logarithmic scale. Tukey's test was performed for the two alleles independently, results represented by letters above each bar (a-j); the same letter indicates no significant difference in titre.

ANOVA demonstrated linkage of ST-289 alleles *fbpA*-17 and *coxA*-15 with *wsp*-11; and ST-285 alleles *fbpA*-62 and *coxA*-84 with *wsp*-661; this was statistically supported in *B. decurtans* (ID85), *B. neohumeralis* (ID221), *B. bryoniae* (ID157) and *D. axanus* (ID88) (Figure 3.2; Appendix B: Table B.5). Linkage of *wsp* and *fbpA* loci was supported in *B. strigifinis* (ID269) and *B. neohumeralis* (ID248). These results indicated that there was no recombination between the two *Wolbachia* strains in the tested regions of the *Wolbachia* genome and, therefore, there was support that the same two strains of *Wolbachia* were found in six Australian fruit fly species (Figure 3.3).

### 3.4 Discussion

Our study revealed a high incidence of shared *Wolbachia* strains in tropical tephritid fruit flies of Australia, the highest incidence for any studied arthropod host community so far, suggesting frequent horizontal transmission of *Wolbachia* in this tropical fruit fly community. *Wolbachia* MLST and *wsp* sequences were analysed in the eight *Wolbachia*-infected Australian tephritid fruit fly species. We found that two identical *Wolbachia* strains were present as double infections in six of the eight infected species; a seventh species harboured one of those two strains as a single infection. A novel strategy based on relative *wsp* and MLST loci quantification confirmed *Wolbachia* genome integrity across six fruit fly species in which *Wolbachia* strains occurred at widely different titres. The absence of single nucleotide polymorphisms and recombination within these strains suggests that horizontal transmission occurred relatively recently.

*Horizontal transmission of* Wolbachia *between Australian tephritid species*

Horizontal transmission of *Wolbachia* between distantly related hosts can occur in three ways (Raychoudhury *et al.* 2009): sharing of the same ecological niche (Rigaud and Juchault 1995), via trophic interactions such as with parasitoids, parasites or predators (Werren *et al.* 1995b), or through hybridisation and introgression (Rousset and Solignac 1995). Tephritid species, particularly in tropical Australia, are somewhat distinctive in their extensive sharing of both host fruits and

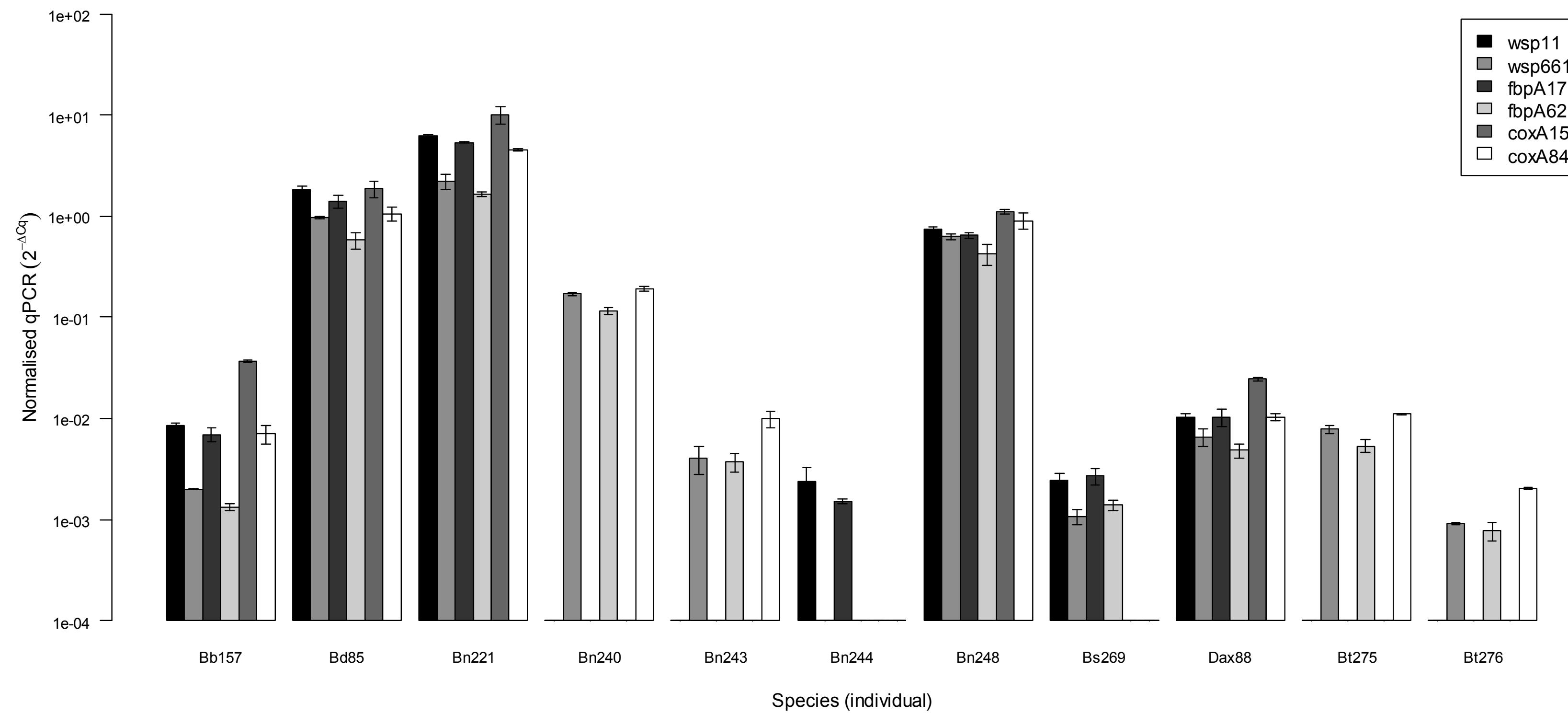



**Figure 3.2** Relative titre of *wsp*, *fpbA* and *coxA* alleles shared between singly and doubly-infected individuals of six tephritid species. Quantification cycle values (Cq) were normalised (see Appendix B: Table B.4) and graphed on a logarithmic scale. Alleles linked on each *Wolbachia* chromosome (*wsp*-11, *fpbA*-17 and *coxA*-15 on strain "1" and *wsp*-661, *fpbA*-62 and *coxA*-84 on strain "2", see Figure 3.3) within an individual have similar titres (see Table B.5).

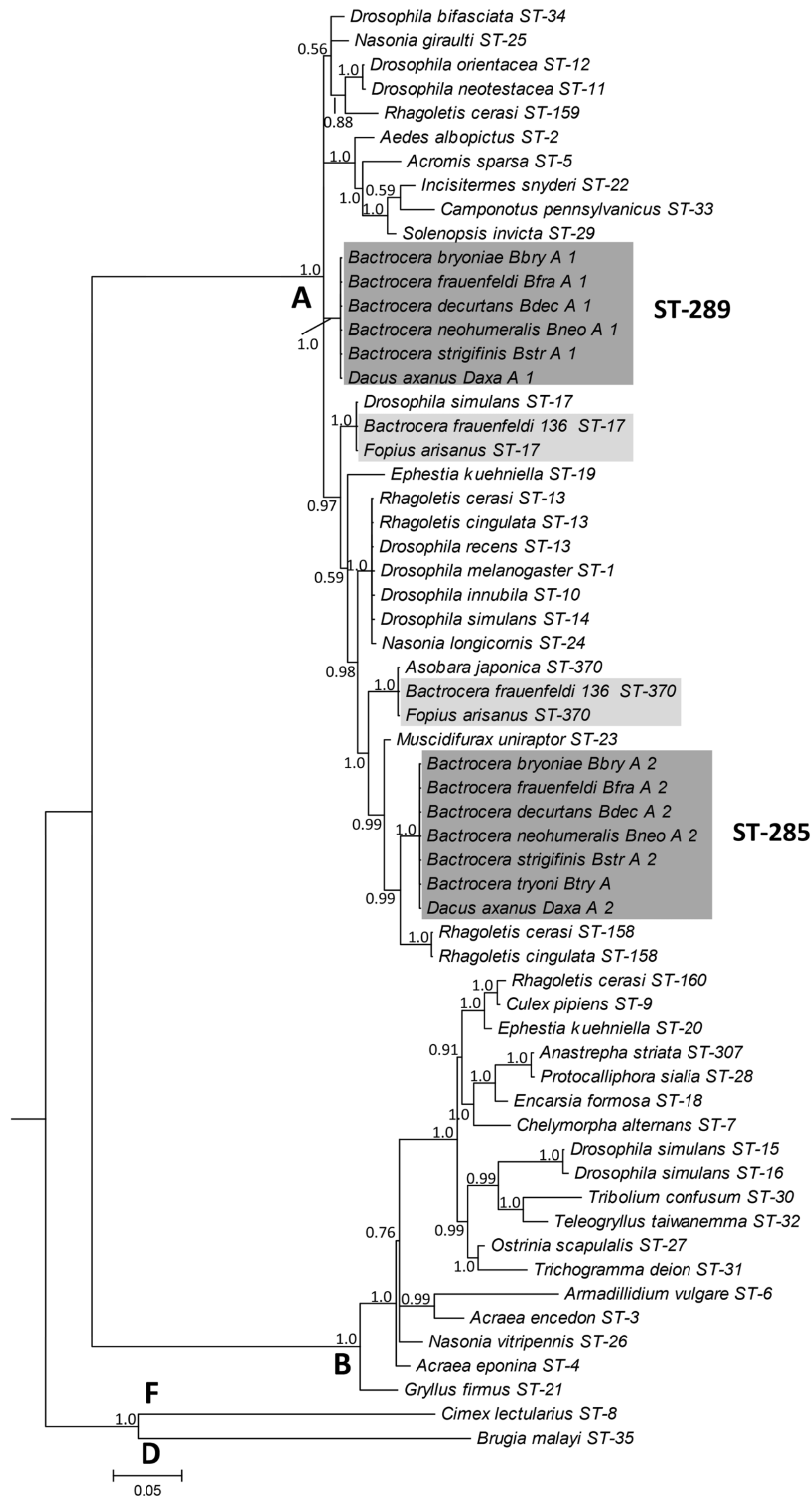



**Figure 3.3** Bayesian inference phylogenetic tree derived from concatenated MLST DNA sequences from this study (highlighted) and selected sequence types from the *Wolbachia* MLST database. Dark grey boxes highlight two *Wolbachia* strains found in six tephritid species (including double infections in six species, and single infection in *Bactrocera tryoni*). Light grey boxes highlight candidate strains found as double infections in *Bactrocera frauenfeldi* and *Fopius arisanus*. Numbers at nodes represent posterior probabilities >50% and letters signify *Wolbachia* supergroups.

parasitoids. For example, *B. neohumeralis* and *B. tryoni* are polyphagous and have been bred from 44 and 48 families of host plants, respectively (Hancock *et al.* 2000). Both species overlap geographically and ecologically with each other and with other species, including *B. frauenfeldi* (21 host plant families), *B. bryoniae* (5 host plant families) and *B. strigifinis*, *B. decurtans* and *D. axanus* (1 host plant family). Uptake of *Wolbachia* by developing larvae within shared fruits via ingestion or injuries may provide opportunities for *Wolbachia* transfer between species. In isopods, *Wolbachia* can be transferred through hemolymph contact (Rigaud and Juchault 1995); *Wolbachia* transfer may also occur through feeding on the same plant (Sintupachee *et al.* 2006) or tissue (Huigens *et al.* 2000). It has also recently been demonstrated that *Rickettsia* symbionts (Alphaproteobacteria) can be transmitted between host insects via plant tissue (Caspi-Fluger *et al.* 2012).

Parasitoids from the genera *Diachasmimorpha* and *Fopius* have multiple tephritid hosts and are found along much of eastern Australia (Carmichael *et al.* 2005). We detected two different *Wolbachia* strains in the parasitoid *F. arisanus* that shared alleles with two *Wolbachia* strains of *B. frauenfeldi* (ID136), suggesting that both hosts are infected by the same *Wolbachia* strains. One *B. cacuminata* larva also carried identical *wsp* alleles, also suggesting that both hosts are infected by the same *Wolbachia* strain. However, it cannot be ruled out that both individuals, one adult *B. frauenfeldi* and one larval *B. cacuminata,* were exposed to these two *Wolbachia* strains or their DNA in the field, e.g. through parasitism by an infected wasp (although wasp DNA was not detected in the larva).

*Fopius arisanus* is native to the Asia-Pacific, introduced to Australia in 1956 (Snowball *et al.* 1962), and is not a demonstrated parasitoid of *B. frauenfeldi.* However, it is a known parasitoid of 17 species of *Bactrocera*, including *B. cacuminata*, and other *Fopius* species are known to parasitise *B. frauenfeldi* (Carmichael *et al.* 2005). *Bactrocera frauenfeldi* has been found widely in the South Pacific and was introduced into northern Australia in 1974 (Drew *et al.* 1978). Opportunity for transfer of *Wolbachia* from parasitoid to host or vice versa exists, and this theoretical capacity has previously been demonstrated in laboratory experiments (Heath *et al.* 1999, Huigens *et al.* 2004) but rarely proven in the field by using the full characterisation of MLST markers.

Furthermore, *Wolbachia* may be vectored between tephritids of different infection status by parasitoids through sequential oviposition. While we cannot address this latter avenue of horizontal transmission, we have an indication for the former in the sharing of *Wolbachia* strains between parasitoid *F. arisanus* and tephritids *B. frauenfeldi* and *B. cacuminata*. *Fopius arisanus* individuals were sampled from Richmond, New South Wales, a temperate area where potential fruit fly hosts such as *B. tryoni*, *B. cacuminata* and *Dirioxa pornia* were all uninfected (Chapter 2); and from Atherton, Queensland, within the tropical range of several infected fruit fly species, including *B. frauenfeldi*. All *F. arisanus* individuals carried *Wolbachia*, evidence of an inherited infection in this parasitoid species, but only one adult fruit fly of 34 *B. frauenfeldi* and one larva of 31 *B. cacuminata* carried this same combination of alleles. Further samples are needed to test whether these are established and inherited infections in the two *Bactrocera* species; alternatively, our findings are evidence for field exposure of these two fruit fly species to *Wolbachia* from their parasitoid, and thus demonstrate an opportunity for horizontal transfer that may result in the colonisation of new host species.

Our findings indicate, based on the higher *Wolbachia* prevalence in the parasitoid, a directionality of *Wolbachia* transfer from parasitoid to host species, although one may expect such a directionality to be less likely given that it would require events of failed parasitism or transfer from ovipositor during the search for the host egg. Overall, our study suggests that sharing of *Wolbachia* between fruit fly species (within same family) may be more common than between fruit fly species and their parasitoids (across different orders), however a more intensive screening of other parasitoids of Australian fruit flies will be required to determine if other parasitoids share similar or identical *Wolbachia*, and if the extent of sharing compares to the high incidence of identical *Wolbachia* strains in their tephritid host species.

A third possible route of horizontal transmission across more closely-related species is through hybridisation and introgression. The ability of *B. tryoni* and *B. neohumeralis* to produce viable offspring under laboratory conditions is well known, and hybridisation in the field is likely to occur at low rates (An *et al.* 2002, Morrow *et al.* 2000, Wang *et al.* 2003). *Bactrocera jarvisi* and *B. tryoni*, a more distantly-related pair of species, can also hybridise in the laboratory (Cruickshank *et al.* 2001, Shearman *et al.* 2010), but little is known of the extent of any hybridisation

between these two and other *Bactrocera* species in the field. However, the presence of the same two independent *Wolbachia* strains in six species belonging to two genera cannot readily be explained this way, particularly across the division between *Bactrocera* and *Dacus*. Our previous analysis of the mitochondrial *cytochrome oxidase subunit I* (*COI*) also excluded the incidence of any shared mitochondrial haplotypes between species except for *B. tryoni* and *B. neohumeralis* (Morrow *et al.* 2000; Chapter 2) as would be expected in the instance of species hybridisation.

For transmission of *Wolbachia* across species boundaries to result in stable establishment of inherited infections, it must be accompanied by adaptation to the new host. Artificial transfer experiments show the increased difficulty of stable inheritance of *Wolbachia* transferred to distantly-related hosts (McMeniman *et al.* 2008). There is also evidence for strain-dependent success of infections establishing in new hosts, a scenario observed following the artificial transfer of *Wolbachia* from one multiply-infected species into an uninfected recipient species (Riegler *et al.* 2004, Zabalou *et al.* 2004). Phylogenetically, *Bactrocera* and *Dacus* may have split approximately 80 million years ago, more recently than the divergence of *Bactrocera*, *Ceratitis* and *Rhagoletis*, and of tephritids from drosophilids (Krosch *et al.* 2012), groups for which three experimental transfers have been successfully maintained over many generations (Apostolaki *et al.* 2011, Riegler *et al.* 2004, Zabalou *et al.* 2004). For vertical transmission to occur, *Wolbachia*, once in individuals of a new host species, must be localised at sufficient density in the germ line (Boyle *et al.* 1993). Ultimately, for *Wolbachia* to perpetuate in host populations, it must confer a selective advantage (eg. Hedges *et al.* 2008), or induce a reproductive phenotype such as CI (Hoffmann *et al.* 1990) without causing too many other deleterious fitness costs, such as through high bacterial densities (Le Clec'h *et al.* 2012). We also detected *Wolbachia* pseudogenes in two species, *B. peninsularis* and *B. perkinsi*, that may be due to lateral gene transfer from *Wolbachia* into the host genomes followed by accumulation of mutations (Dunning Hotopp *et al.* 2007). It will be interesting to follow this up with an RT-PCR approach on RNA extracts of fresh field material. Our genomic extracts were performed on samples collected in field traps between 1997 and 2012 and thus not ideal for a transcription analysis. *Wolbachia* pseudogenes in host genomes are interesting from the perspective that

they could represent footprints of historic *Wolbachia* infections that may have gone extinct in these host lineages.

*Linking* Wolbachia *alleles in multiply-infected individuals using quantitative PCR*

In our study we established quantitative PCR as a new method to assign and link alleles of two A supergroup *Wolbachia* genomes that were harboured in individual hosts as double infections. While both strains were found singly in *B. neohumeralis*, only double infections were found in four other species. It cannot be assumed that finding identical *Wolbachia* alleles necessarily indicates they are part of the same strains when present across a number of host species, because of high levels of recombination between *Wolbachia* strains (Baldo *et al.* 2006a). Numerous examples exist of *Wolbachia* strains with the same *wsp* allele, but different alleles at other loci, or vice versa (Baldo and Werren 2007, Baldo *et al.* 2008, Doudoumis *et al.* 2012, Guidolin and Cônsoli 2013, Salunke *et al.* 2012). Consequently, Baldo *et al.* (2006b) established that at least three genes are necessary to type strains accurately, and this can serve as an indicator of similarity at other chromosomal genes. In the case of double infections, determining linkage of three or more loci by quantitative measurements may be sufficient to discount intergenic recombination. We successfully applied this criterion to our study system because the supergroup A allele sequences were suitable for highly specific qPCR primer design, and titres of the two strains were not equal.

A discernible difference in overall *Wolbachia* titre between individuals was found. Within *B. neohumeralis*, strain density within the abdomen varied more than 7,000-fold for *wsp*-11 and more than 1,600-fold for *wsp*-661 over the eight examined individuals. Unckless *et al.* (2009) detected a 20,000-fold difference in *Wolbachia* titre among field caught *Drosophila innubila* female ovaries, and Ahantarig *et al.* (2008) found *Wolbachia w*AlbA varied by 180,000-fold in whole *Aedes albopictus* field specimens. These wide differences in *Wolbachia* density may be due to ecological effects on vertical transmission rates and may beget the overall low prevalence of *Wolbachia* that we previously detected in *Bactrocera* and *Dacus* species (Chapter 2). However, it must also be considered that all specimens in our study were field-caught and therefore of unknown age; different host ages can

account for changes in *Wolbachia* titre (Arthofer *et al.* 2009b). Furthermore, all flies included in this study were males, due to the collection methods using male attractants, which could also result in an underestimation of *Wolbachia* prevalence (Chapter 2).

The relative density of the two *Wolbachia* strains within five doubly-infected *B. neohumeralis* males was modest (0.6 to 2.8) in light of the overall titre differences between strains found in other host species, such as a more than 500,000-fold difference in A and B strain density in some *A. albopictus* individuals (Ahantarig *et al.* 2008). Mouton *et al*. (2003) found the density of three *Wolbachia* strains in a single *Leptopilina* parasitoid wasp to be independent from each other, whereas Watanabe *et al*. (2011) discovered that overall *Wolbachia* titre was regulated in an *Orius* bug species and that one *Wolbachia* suppressed another within a superinfection. In our study, the small number of analysed individuals makes predictions of synergistic interactions difficult, but neither strain showed evidence of suppression at low or high densities, and overall density did not seem to influence the presence or absence of a given strain. While it was more common to find ST-289 at higher densities than ST-285, still two of the six superinfected *B. neohumeralis* individuals carried ST-285 at slightly higher densities and both strains were present as single infections.

### 3.5 Conclusions

Despite ample indirect (evolutionary) and some direct (ecological) lines of evidence for horizontal transmission of *Wolbachia* between host taxa, the majority of studies so far have been based on a single variable *Wolbachia* marker gene (Haine *et al.* 2005, Kittayapong *et al.* 2003, Sintupachee *et al.* 2006). A small number of more recent studies have used MLST to better identify the pandemic of *Wolbachia* across different host taxa, although so far just within the same trophic level. These studies have revealed incidences for horizontal transmission for a small number of host species with *Wolbachia* strains that are highly similar (Baldo *et al.* 2008, Raychoudhury *et al.* 2009, Salunke *et al.* 2012, Stahlhut *et al.* 2010) or identical (Schuler *et al.* 2013). Our study demonstrates for the first time an overall high incidence of identical and multiple *Wolbachia* strains within the community of

Australian tropical tephritid fruit flies as well as their parasitoids. These fruit flies share food resources and a diverse parasitoid complex; both may facilitate the horizontal transmission of *Wolbachia*, although it remains to be analysed in what relative proportions. The sharing of the same two *Wolbachia* strains across six species of the genera *Bactrocera* and *Dacus*, the sharing of a single infection between two closely related sibling species that can hybridise, as well as the identity of *Wolbachia* strains between a tephritid species and a parasitoid species of tephritids, all present opportunities to further investigate the ecological and evolutionary dimensions in the horizontal transmission of *Wolbachia* between host species. This may be achieved through more extensive field studies that target both horizontal and vertical transmission of *Wolbachia* in this insect community as well as the characterisation of highly polymorphic markers such as variable number tandem repeats in the *Wolbachia* genome (Riegler *et al.* 2012). Combined with analysis of the frequency of lateral gene transfer of *Wolbachia* into host genomes and extinction rates of *Wolbachia* infection in host lineages, this will enable investigation into the life-history cycle of *Wolbachia* infections (Schneider *et al.* 2012).

## 3.6 Experimental procedures

*Insect samples*

A total of 104 male fly whole genomic DNA extracts were used in this study, from eleven *Bactrocera* and three *Dacus* species, comprising a subset of the 24 fruit fly species previously sampled and analysed (Chapter 2). Samples were selected based on prior identification of *Wolbachia wsp* or 16S rDNA (44 samples; Appendix B: Table B.1) and additional uninfected samples primarily from species with other individuals harbouring *Wolbachia* (60 samples; Appendix B: Table B.1). Specimens were collected in northern Queensland during summer, except for uninfected control individuals of *B. cacuminata*, *B. neohumeralis, B. tryoni, D. aequalis* and *D. newmani* from southeast Queensland, New South Wales and the Northern Territory (Appendix B: Table B.1). Flies were collected by using traps equipped with male attractant cue lure (Osborne *et al*., 1997; Royer and Hancock, 2012) except for individuals of *B. cacuminata* that were collected from infested wild tobacco fruits in Richmond, New South Wales. Besides adult fruit fly individuals we included ten

*B. cacuminata* larvae collected from wild tobacco fruits in Atherton, Queensland (Table 3.3), seven parasitoid wasps that had emerged from pupated *B. cacuminata* of the same population in Queensland, and seventeen wasp parasitoids collected from an orange orchard in Richmond, New South Wales (Table 3.2). The wild tobacco fruit fly larvae and adult parasitoids were given unique identification numbers, and entire specimens were used for DNA extraction.

### *Polymerase Chain Reaction*

Primers for five MLST loci: *gatB, coxA, hcpA, ftsZ* and *fbpA* (Baldo *et al.* 2006b), were used to resolve relationships between *Wolbachia* found in the different host species. PCR protocols are detailed in Table B.6 (Appendix B). MLST standard primer sets; external primer sets (denoted F3 and R3) used in nested PCR with standard primers for *gatB*, *hcpA* and *fbpA*; and B-group specific primer pairs were used as specified in the *Wolbachia* MLST database (http://pubmlst.org/wolbachia/; Jolley *et al.* 2004). *Wolbachia* allele-specific end-point PCR employed primers designed in this study for qPCR (see below; Appendix B: Table B.2).

DNA barcoding of the parasitoids using primers LCO1490 and HCO2198 (Folmer *et al.* 1994) was used to resolve species identity. PCR conditions (Rugman-Jones *et al.* 2007) are described in Table B.6 (Appendix B). All fruit fly larvae *COI* amplicons (primers Dick and Pat; Simon *et al.* 1994) were direct sequenced and compared to sequences generated in Chapter 2 for species identification. Measures to minimise potential contamination included treatment of all individuals with 4% sodium hypochlorite (Sigma, St Louis, MO) prior to extraction, use of filter tips, regular replacement of stock solutions as well as fresh aliquots for individual experiments, plus independent DNA extraction and PCR experiments in different laboratories (Table 3.2, Table 3.3).

### *Cloning and sequencing*

Cloning of amplicons and sequencing of 3 to 5 clones was performed for each MLST locus for one specimen of each species, except for singly-infected *B. tryoni*. Direct

sequencing of each amplified MLST locus for each species, including doubly-infected species, was also performed to obtain sequences from single infections in *B. tryoni*, *B. neohumeralis* and *B. frauenfeldi*; as well as to confirm multiple infections and polymorphic sites.

For cloning, 3µL PCR product was used directly in the ligation reaction, with 25ng pGEM-T Easy vector (Promega), 1X Rapid ligation buffer and 3u T4 DNA ligase (Promega), incubated at 4°C overnight. Transformation of JM109 competent cells (Promega) was according to manufacturer's protocol. Clones were PCR screened by smearing the colony into a PCR tube using a sterile pipette tip, and assembling reactions according to Table B.6 (Appendix B); at least three positive clones, recognised by appropriately sized PCR products, were sequenced. Preparation of PCR amplicons for direct sequencing involved treatment with a combination of 0.5u Exonuclease I (New England Biolabs, Ipswich, MA) and 0.25u Shrimp Alkaline Phosphatase (Promega), with incubation at 37°C for 30min, then 95°C for 5min. Sequencing of each amplicon, in both directions using the primers used in the PCR, was carried out by Macrogen (Seoul, Korea).

*Quantitative PCR*

Individuals of five fruit fly species (*B. bryoniae*, *B. decurtans*, *B. neohumeralis*, *B. strigifinis* and *D. axanus*) produced the same allele sequence variants for two *Wolbachia* strains, suggesting that they may be infected by the same strains. These two *Wolbachia* strains were then tested for absence of chromosomal recombination by allele-specific qPCR. Successful application relied upon the allele sequences from each *Wolbachia* strain being sufficiently diverged to design efficient primers that differentiated the alleles in an amplification reaction. DNA sequences for *wsp*, *coxA* and *fbpA* met this criterion: the two alleles at each locus contained 61, 12 and 18 SNPs, respectively. The alleles for *gatB*, *ftsZ* and *hcpA* contained 1, 5 and 2 SNPs and were thus not included in this analysis. Primer pairs were designed to highly polymorphic regions of *wsp*, *coxA* and *fbpA* using the primer-design tool PrimerQuest (www.idtdna.com/Scitools/Applications/Primerquest/Default.aspx; Appendix B: Table B.2). The single copy gene *scarlet* was chosen as reference gene (Zhao *et al.* 1998) for which *Dacus*-specific and *Bactrocera*-specific primers were

designed on previously obtained partial sequences for several species of *Bactrocera* and *Dacus* (Curthoys 1997; Appendix B: Table B.2).

Strain-specificity of *Wolbachia* qPCR primers was validated in several ways. Quantitative PCR targeting the single strains in *B. tryoni* (ID275) and *B. neohumeralis* (ID244) at the three loci showed no evidence of amplification of the alternative alleles. PCR-RFLP, using TaqI (Promega) restriction enzyme, according to the manufacturer's instructions, with an incubation period of 3h, was applied to *wsp* amplicons; when visualised by gel electrophoresis the product for *wsp*-11 was uncut while the *wsp*-661 product digested completely into two fragments, 70bp and 32bp in length. Melting curves for each primer pair generated after qPCR by an incremental temperature increase, showed a single defined peak for each primer set, representing a single amplicon disassociating at a temperature characteristic for that sequence, thus validating the specificity of the primers designed to each allele.

Efficiency of *Wolbachia* primer pairs was assessed by qPCR, directed at three separate template types. Firstly, amplicons were generated by amplifying *B. neohumeralis* (ID248) DNA (harbouring both *Wolbachia* strains) using all six primer pairs separately, and the amplicons gel-extracted using the Wizard SV Gel and PCR Clean-up System (Promega) according to the manufacturer's instructions. Serial dilutions of these amplicons, *B. decurtans* (ID85) genomic DNA (also harbouring both strains), and a third template comprising a mixture of equal volumes of genomic DNA from individuals from four of the species used in the experiment – *B. neohumeralis* (ID109), *B. decurtans* (ID85), *B. strigifinis* (ID269) and *D. axanus* (ID88) – were subjected to triplicate qPCR. Similarly, the efficiency of *Dacus*-specific *scarlet* primers was tested on *D. axanus* (ID88) genomic DNA, and *Bactrocera*-specific *scarlet* primers were individually tested on *B. neohumeralis* (ID248), *B. decurtans* (ID85), *B. strigifinis* (ID269) and *B. bryoniae* (ID157). Efficiency (E) of the primers was examined by provision of a standard curve using the Rotor-gene 6000 software (version 1.7); all primer efficiency values were greater than 0.95 (Appendix B: Table B.7). The primers were therefore deemed suitable for qPCR analysis using the Delta Cq model without efficiency correction (Schmittgen and Livak 2008).

All sample and standards qPCRs were assembled in triplicate using a CAS-1200 pipetting robot (Corbett Research) in 100 well rotor-discs. Reaction volumes were 10µL and comprised 5µL SensiMix SYBR (Bioline, Sydney, Australia), 0.4µM each primer and 4.2µL DNA template. Each DNA sample was diluted 1/10 prior to addition to the reactions. Cycling conditions were 95°C for 10min, then 45 cycles of 95°C for 20s, 62°C for 30s and 72°C for 45s, followed by a disassociation cycle with incremental increase of 1°C from 60°C to 98°C every 5s. Quantification cycle (Cq) values for reference and target genes were calculated from the average replicate Cq values at the same threshold level. *Wolbachia* Cq values were normalised to *scarlet* as individual data points using $2^{-\Delta Cq}$ (Schmittgen and Livak 2008).

*Statistical analyses*

Statistical analyses (ANOVA and Tukey's test) were performed using R 2.15 (R Core Team 2012). In order to perform allele assignment based on qPCR, the alleles of loci within one *Wolbachia* genome were grouped, as their quantified values related directly to the relative density of the two different *Wolbachia* chromosomes, and tested by ANOVA. Conversely, grouping the wrong combination of alleles should return a non-significant result. In order to test the likelihood of independently deriving the same MLST alleles and ST from a host (*B. frauenfeldi*) and its parasitoid species (*F. arisanus*) we performed binomial sampling of the MSLT database (accessed in December 2013).

*Phylogenetic analyses*

DNA sequences were trimmed and edited in Sequencher 4.0 (GeneCodes Corp), and deposited in GenBank (accession numbers KC668333-KC668409; KC581413-KC581417; KC857545-KC857548) and the *Wolbachia* MLST Database (http://pubMLST.org/wolbachia/). Sequence alignment utilised the MUSCLE algorithm within Mega 5.05 (Tamura *et al.* 2011). Five MLST genes were concatenated and aligned with those of the following strains from the *Wolbachia* MLST database: ST-1 to ST-35 (Baldo *et al.* 2006b); tephritid fruit fly ST-13, ST-158 to ST-160 (Arthofer *et al.* 2011) and ST-307 (Martínez *et al.* 2012). The double

infections in *F. arisanus* and *B. frauenfeldi* (ID136) were not tested for allele linkage by qPCR, but a potential combination of alleles based on the existing *w*Ri (ST-17) and *A. japonica* (ST-370) strains was used. MrBayes 3.2 (Ronquist *et al.* 2012) was employed to calculate Bayesian phylogenetic inferences, with evolutionary model GTR+G+I selected from the Find Best Fit DNA Model in Mega 5.05. Parameters selected for the run were a maximum of $10^7$ generations, stopping earlier if convergence is reached, sampling every 100 generations, with the first 25% of trees discarded and 50% majority rule tree returned.

# Chapter 4

# The microbiome of field-caught and laboratory-adapted Australian tephritid fruit fly species with different host plant use and specialisation

Morrow, J. L., Frommer, M., Shearman, D.C.A. and Riegler, M. "The microbiome of field-caught and laboratory-adapted Australian tephritid fruit fly species" Microbial Ecology (in preparation).

## 4.1 Abstract

Tephritid fruit fly species display a diversity of host plant use on a scale from monophagy to polyphagy. Furthermore, while some fruit fly species prefer ripening fruit others are restricted to damaged or rotting fruit. Such a diversity of specialisation may be reflected in the microbial symbiont diversity of tephritids and their grade of dependency on their microbiomes. Here we investigated the microbiome of six tephritid fruit fly species from three genera, including species that are polyphagous pests (*Bactrocera tryoni, Bactrocera neohumeralis, Bactrocera jarvisi*, *Ceratitis capitata*) and a monophagous specialist (*Bactrocera cacuminata*). These were compared with the microbiome of a non-pestiferous but polyphagous tephritid species that is restricted to damaged or rotting fruit (*Dirioxa pornia*). The bacterial community associated with whole fruit flies was analysed by 454 pyrosequencing of 16S rDNA amplicons. Overall, the dominant bacterial families were Enterobacteriaceae and Acetobacteraceae (Proteobacteria), and Streptococcaceae and Enterococcaceae (Firmicutes). Comparisons across species and genera found different microbial composition in the three genera, but limited consistent differentiation between *Bactrocera* species. Comparisons of field-collected versus laboratory-reared flies exposed streamlining of microbial diversity as a consequence of laboratory adaptation. Overall, more diversity was found in polyphagous species over the monophagous species. The microbiome of *D. pornia* was most distinct from the other five tephritid fruit fly species, which may be due to its ecologically different niche of rotting or damaged fruit, as opposed to ripening fruit favoured by the other species. Our study is the first amplicon pyrosequencing study to compare the microbiomes of tephritid fruit fly species and thus delivers important information about the turnover of microbial diversity within and between fruit fly species and their potential application in pest management strategies.

## 4.2 Introduction

Insect species are associated with a vast array of symbiotic microorganisms, and outcomes of these symbioses for hosts range from beneficial to deleterious. For example, plant-sap feeding aphids benefit nutritionally from associations with primary obligate endosymbionts (Prosser and Douglas 1991) while they can receive parasite protection from facultative secondary endosymbionts (Oliver *et al.* 2003); wood feeding termites rely heavily on gut microbes including protists and bacteria for nutritional provisioning (Hongoh *et al.* 2008, Warnecke *et al.* 2007); furthermore, insect symbionts can manipulate reproduction and increase or decrease fecundity (Werren *et al.* 2008), influence mating preference (Sharon *et al.* 2010), and cause or mitigate pathogenicity (Ceuppens *et al.* 2013, Scarborough *et al.* 2005).

The important role of microbial symbionts in the life history of tephritid fruit flies was first demonstrated in the olive fly, *Bactrocera oleae*, where removal of its symbionts by antibiotic treatment led to reduced larval growth in its olive host (Hagen 1966) in a diet-dependent fashion (Ben-Yosef *et al.* 2010). The identity of the main symbiont of *B. oleae* was assigned to the unculturable *Candidatus* Erwinia dacicola (Capuzzo *et al.* 2005) while another study also proposed culturable *Acetobacter tropicalis* as a major symbiont (Kounatidis *et al.* 2009). *Candidatus* E. dacicola is an abundant bacterium located in the oesophageal bulb, midgut and ovipositor of field individuals (Capuzzo *et al.* 2005, Estes *et al.* 2009). Its role as an obligate symbiont was originally demonstrated by localisation to a specialised storage structure, the oesophageal bulb (Petri 1910), then through vertical transmission, and nutritional supplementation of the host (Capuzzo *et al.* 2005). A co-evolutionary interaction was suggested for tephritids of the subfamily Tephritinae, that develop in the flower heads of Asteracea, and their symbiont *Candidatus* Stammerula tephritidis (Mazzon et al. 2008, Mazzon et al. 2010). However in other tephritids, such as *Rhagoletis* spp., no such consistent association was found with any particularly obligate symbionts (Howard *et al.* 1985), while nutritional benefits through the presence of symbiotic bacteria supported increased fecundity (Tsiropoulos 1981). Similarly, observed in the Mediterranean fruit fly, *Ceratitis capitata*, are increased longevity (Behar *et al.* 2008b) and symbiont-driven nitrogen fixation by Enterobacteriaceae, providing a potential source of protein (Behar *et al.* 2005). Symbiont-facilitated nitrogen fixation was also observed in Queensland fruit

fly *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae), however its role as protein source is unknown (Murphy *et al.* 1988, Murphy *et al.* 1994).

Our study included the microbiomes of flies from six tephritid species, whereby the focal points sat with four species of *Bactrocera* (subfamily Dacinae) that were compared to *C. capitata* (subfamily Ceratitidinae) and *Dirioxia pornia* (subfamily Trypetinae). *Bactrocera tryoni*, *B. neohumeralis*, *B. jarvisi* and *B. cacuminata* are four fruit fly species useful for inter-specific comparisons, as they display a spectrum of behaviours, such as different scales of host plant specialisation, varying climatic tolerance, and attraction to different male lures (Hancock *et al.* 2000). *Bactrocera tryoni* and *B. neohumeralis*, the common and lesser Queensland fruit fly, respectively, are sympatric, highly polyphagous pest species found extensively in eastern Australia, but unlike *B. tryoni*, *B. neohumeralis* is not found in southerly temperate areas (Osborne *et al.* 1997). Wild tobacco fruit fly, *B. cacuminata*, and Jarvis' fruit fly, *B. jarvisi*, are both widespread from tropical to temperate regions of eastern and north-eastern Australia, respectively, but *B. cacuminata* is essentially a monophagous specialist on invasive *Solanum mauritianum* (wild tobacco) and perhaps a few other solanaceous fruits, while *B. jarvisi* utilises host species from a number of plant families but exhibits a particular preference for native *Planchonia careya* (cocky apple; Lecythidaceae) (Hancock *et al.* 2000). Mediterranean fruit fly, *C. capitata*, was accidentally introduced and established in both New South Wales and Western Australia in the late 19$^{th}$ century but disappeared from the east coast in the mid-20$^{th}$ century (Dominiak and Daniels 2012). In comparison to these fruit flies that mostly infest ripening but undamaged fruit, Island fruit fly, *D. pornia*, widespread in Australia, is not a pest as it is restricted to infesting damaged fruit from a wide variety of plant families.

Symbiotic bacteria of tephritid fruit flies may play an important role in host nutrition, development, fecundity, climatic adaptation, as well as behaviour such as host plant use and preference, thus influencing pest status. Moreover, a detailed characterisation of the tephritid microbiome and identification of key bacterial symbionts could potentially improve management strategies that rely on trapping or baiting of field flies (e.g. lure-and-kill methods; Dominiak and Ekman 2013) or on the fitness of mass-reared sterile or incompatible flies used for the sterile insect technique (SIT;

Ben Ami *et al.* 2010, Sacchetti *et al.* 2014) and the incompatible insect technique (IIT; Boller *et al.* 1976, Zabalou *et al.* 2009).

Until quite recently, classical microbiological methods such as culture-dependent isolation and morphological and physiological characterisation have been used extensively to identify microorganisms associated with tephritid fruit flies. Various adult tissues were examined in this way, including the gut or sections of the gut (Drew and Lloyd 1987, Fitt and O'Brien 1985, Marchini *et al.* 2002, Murphy *et al.* 1994, Thaochan *et al.* 2010), oesophageal bulb (Marchini *et al.* 2002), head (Fitt and O'Brien 1985) and other life stages such as eggs and pupae (Fitt and O'Brien 1985). More recently, molecular methods have been used to identify both culturable and non-culturable bacteria, primarily targeting the 16S rDNA through PCR amplification, and cloning and sequencing (Capuzzo *et al.* 2005, Thaochan *et al.* 2010), denaturing gradient gel electrophoresis (DGGE; Behar *et al.* 2008a, Behar *et al.* 2008b) or next-generation sequencing (Aharon *et al.* 2013).

Microbial diversity and abundance are influenced broadly by environmental or species-specific factors. Here we used 454 pyrosequencing of the 16S rRNA gene to comparatively survey the microbiomes of both field and laboratory flies of six species of the genera *Bactrocera*, *Ceratitis* and *Dirioxa*. Deep sequencing has the power to reveal the presence of rare and unculturable bacteria, and has so far only been applied once to the analysis of the microbiome of a tephritid fruit fly, *C. captitata* (Aharon *et al.* 2013). Our analysis extended previous work on the microbiome of *Bactrocera* and *Ceratitis* species, which had identified Enterobacteriaceae and Pseudomonaceae by culture-based methods (Fitt and O'Brien 1985), as well as Leuconostocaceae, Enterococcaceae and Acetobacteraceae by 16S rDNA cloning methods (Behar *et al.* 2005, Behar *et al.* 2008b, Thaochan *et al.* 2010). We also included data on the microbiome of *D. pornia*, a species with an extensive geographic and host fruit range, that inhabits a different ecological niche of damaged or rotting fruit.

The aim was to canvas for the first time the entire microbiome of selected Australian tephritid fruit fly species, using a deep sequencing approach, to discern low abundance taxa and confirm highly abundant taxa. We envisaged expanding on previous culture-based work that would have failed to detect unculturable bacteria,

and on culture-independent methods of 16S rDNA cloning and sequencing, that would have failed to detect rare bacterial taxa. We further used the opportunity of parallel amplicon sequencing to compare microbial diversity within Australian species of the genus *Bactrocera*, and between *Bactrocera* and other genera, to test host-phylogenetic constraints of microbiome composition. We also contrasted field-collected with laboratory-reared flies, and species with different scales of host plant specialisation (polyphagous versus monophagous) and host plant use (undamaged versus damaged and fermenting fruit), with an expectation of reduced diversity in the monophagous species and more significant microbiome differences in species with different host plant use.

## 4.3 Materials and methods

*Sample collection*

Adult flies of natural populations of *B. tryoni* (BtF12) and *D. pornia* (DpF12) were collected in January 2012 from various plants in the orchard on the Hawkesbury campus (Richmond) of University of Western Sydney, NSW. In 2009, wild tobacco fruits infested with *B. cacuminata* (BcF09) larvae collected in Summer Hill, NSW, were brought into the laboratory and emerging adults collected. Laboratory lines of *B. tryoni* (BtGWS09, BtHWS09, Bt54WS10), *B. neohumeralis* (BnWS09, BnQD09, BnQD12), *B. cacuminata* (BcWS09), *B. jarvisi* (BjWS10) and *C. capitata* (CcWA99) were established in either 1999, 2009, 2010 or 2012, and maintained over various generations in three different laboratories, at UWS, Queensland Department for Agriculture, Fisheries and Forestry (QDAFF) or Department of Agriculture and Food Western Australia (DAFWA) (Appendix C: Table C.1). All UWS laboratory stocks were maintained under the same environmental conditions, on larval diet of dried, shredded carrot (133g/L), methylparaben (2.78g/L) and Torula deactivated yeast (55.6g/L) made into a slurry with hot tap water. The two *B. neohumeralis* stock lines from QDAFF were raised on a similar diet with different concentrations of dried diced carrot (150g/L), methylparaben (3.33g/L) and Torula deactivated yeast (50g/L). Adults from both laboratories were fed sugar and water and a 10:1 protein mix of yeast hydrolysate and sugar (Meats 1981). Medfly lines were raised on larval diet consisting of sucrose (120g/L), brewer's yeast (80g/L), wheat bran (240g/L),

methylparaben (2g/L), sodium benzoate (4g/L) and 9mL HCL made up to 1L with water; and adult diet of 3:1 sucrose and yeast hydrolysate mix (Rossler and Koltin 1976). Flies collected from UWS and QDAFF laboratory stocks had been fed with adult diet one week after emergence and were 2-4 weeks old upon collection, flies from DAFWA were of unknown age, while the age of flies collected in the field was undetermined.

*DNA Extraction and 454 Pyrosequencing*

Eight female specimens of each of the twelve populations or laboratory stock lines were surface sterilised in 4% sodium hypochlorite solution and rinsed thoroughly in 0.2% Triton-X and Milli-Q water. DNA was extracted from individual whole flies using QiaAmp DNA Mini kit (Qiagen), including a 2min, 0.4mg RNase A (Sigma) treatment following cell lysis but prior to the purification steps. An equal volume of each DNA sample for each population was pooled without adjustment for minor variations in DNA concentration. Quality of DNA preparations was ascertained by gel electrophoresis, Nanodrop spectrophotometry, and the concentration of double-stranded DNA determined by the Qubit 2.0 Fluorometric assay (Appendix C: Table C.2). 16S rDNA PCR was primed by 341For 5' CCTAYGGGRBGCASCAG 3' and 806Rev 5' GGACTACNNGGGTATCTAAT 3' which include the V3 and V4 regions of the 16S rDNA sequence (Andersen *et al.* 2013, Yu *et al.* 2005). Amplicon libraries for the twelve samples were prepared using FastStart High Fidelity PCR System dNTPack (Roche). Library preparation, multiplexing of the twelve samples with individual multiplexing identifier sequences, and 454 amplicon pyrosequencing was performed in a single run on the Roche GS Junior platform at the Hawkesbury Institute for the Environment Next-Generation Sequencing Facility (UWS, Richmond).

*Sequence analyses*

Analyses were performed using the Ribosomal Database Project tools (RDP ver10; Cole *et al.* 2009). Initial quality control removed primers and eliminated sequences that were smaller than 150bp, had an average quality score below 20 or contained

any ambiguous bases. Chimeric sequences were detected using Decipher (http://decipher.cee.wisc.edu/FindChimeras.html) and removed. Alpha diversity measures of species richness (Chao1), diversity (Shannon) and evenness were calculated in RDP ver10. RDP Classifier was applied to compare sequences from each sample to the RDP database for taxonomic assignment, with 80% bootstrap cut-off (Wang *et al.* 2007). Sequence alignment (Infernal; Nawrocki *et al.* 2009) and clustering (mcClust; Fish *et al.* 2013) were implemented with default parameters. Operational Taxonomic Units (OTUs) were defined at 97% similarity; this threshold was used to display rarefaction curves and for selecting a sequence representative of each cluster by calculating the minimum sum of square distances between sequences within the cluster. Representative sequences from each sample were aligned in Mega version 5.05 (Tamura *et al.* 2011) using the ClustalW algorithm, visually inspected, and a maximum likelihood tree produced (model: Kimura 2-parameter). The resulting phylogenetic tree, and associated abundance data for each representative sequence, was imported into UniFrac to perform phylogenetic distance-based measurements (Lozupone and Knight 2005). Beta diversity across the twelve samples was measured by both unweighted and weighted UniFrac analyses to assess the relationship between presence and absence of taxa (unweighted) in each sample, as well as incorporating abundance of taxa (weighted). Abundance-based Jaccard distances (Chao *et al.* 2006) were calculated within the RDP pipeline; this distance metric also analysed presence / absence and abundance data, however the non-phylogenetic method treated all clustered sequences equally, regardless of nucleotide distance. Visualisation of samples from Jaccard, unweighted UniFrac and weighted UniFrac distances used principal coordinates plots, implemented in R version 2.15.1 (R Core Team 2012) with Euclidean distances specified. Hierarchical clustering of samples and OTU abundance was performed using pheatmap in R version 2.15.1, using clustering threshold of 95% and Euclidean distances. For comparison, the *COI* mtDNA phylogeny of the six fruit fly species, rooted with *D. melanogaster* sequence, was generated using *Cytochrome oxidase I* sequences (see Chapter 2). Additional statistical analyses were also implemented in the R environment.

### 4.4 Results

*Summary of 454 pyrosequencing data*

Twelve tephritid fruit fly libraries, each consisting of a pool of eight individuals, were subjected to 454 pyrosequencing of ~467bp of the bacterial 16S rRNA gene including primer sequences, covering variable regions V3 and V4. Samples comprised four *Bactrocera* species, including laboratory populations from two different sources as well as field specimens; *C. capitata* from another laboratory and *D. pornia* field specimens (Table 4.1). We obtained 119,368 sequencing reads; following quality filtering and chimera removal, 113,157 reads remained (94.8%), ranging from 3,870 to 17,463 reads per sample (Appendix C: Table C.3). Clustering at 97% identity across all fruit fly species produced 142 operational taxonomic units (OTUs), and the largest OTU contained 24,831 sequences that were represented in seven samples. Alpha diversity measurements such as rarefaction curves (Figure 4.1) and the Chao1 estimates of species richness (Table 4.1) indicated that most but not all of the microbial diversity associated with tephritid fruit flies was captured by sequencing coverage. Chao1, Shannon's diversity index (H) and Equitability (E; measuring evenness of each sample) indicate overall low diversity (Table 4.1; Chandler *et al.* 2011). Field-caught *D. pornia* contained the most OTUs at 97% sequence similarity threshold (38 OTUs, Chao1 48.5) and the most diverse bacterial community was associated with the *B. neohumeralis* QDAFF laboratory line that had been maintained in the laboratory for three years (approximately 24 generations). The alpha diversity of field-collected *B. tryoni* and *B. cacuminata* was higher than in the corresponding laboratory lines. Laboratory lines maintained at UWS and QDAFF showed between-laboratory distinction in species richness (Student t-Test: $t = 2.58$, $df = 6$, $p < 0.05$) and Shannon's diversity (Student t-Test: $t = 2.83$, $df = 6$, $p < 0.05$). Overall, monophagous *B. cacuminata* samples contained lower species richness (Welch t-test: $t = 3.36$, $df = 6.07$, $p < 0.05$) and diversity (Welch t-test: $t = 8.22$, $df = 6.01$, $p < 0.05$) than polyphagous *B. tryoni* and *B. neohumeralis* field and laboratory samples.

**Table 4.1.** Alpha diversity metrics of 454 pyrosequencing of twelve samples, calculated at the 97% identity level.

| sample ID | Line | Environment | N | OTUs | Chao1 | H' | E |
|---|---|---|---|---|---|---|---|
| BtGWS09 | *B. tryoni* GOS | UWS laboratory | 3870 | 13 | 20.5 | 0.965 | 0.376 |
| BtHWS09 | *B. tryoni* HAC | UWS laboratory | 8650 | 13 | 14.5 | 1.219 | 0.475 |
| Bt54WS10 | *B. tryoni* 54 | UWS laboratory | 8384 | 22 | 23.67 | 1.104 | 0.357 |
| BtF12 | *B. tryoni* Field | field | 8391 | 29 | 34.25 | 1.376 | 0.409 |
| BnWS09 | *B. neohumeralis* UWS | UWS laboratory | 6873 | 12 | 13 | 1.005 | 0.405 |
| BnQD09 | *B. neohumeralis* 09 | QDAFF laboratory | 9087 | 29 | 36 | 1.942 | 0.577 |
| BnQD12 | *B. neohumeralis* 12 | QDAFF laboratory | 8834 | 22 | 23 | 1.508 | 0.488 |
| BjWS10 | *B. jarvisi* UWS | UWS laboratory | 17463 | 18 | 20.5 | 1.042 | 0.361 |
| BcWS09 | *B. cacuminata* UWS | UWS laboratory | 10678 | 9 | 12 | 0.233 | 0.106 |
| BcF09 | *B. cacuminata* Field | field | 9748 | 11 | 12.5 | 0.241 | 0.100 |
| CcWA99 | *C. capitata* Medfly | DAFWA laboratory | 10998 | 8 | 14 | 0.086 | 0.041 |
| DpF12 | *D. pornia* Field | field | 10281 | 38 | 48.5 | 1.673 | 0.460 |

N=number of 16S rDNA sequences isolated from the host sample following quality filtering and chimera removal; OTUs = operational taxonomic units calculated at 97% similarity; Chao1 = estimate of species richness; H' = Shannon's diversity index; E = Equitability (evenness)

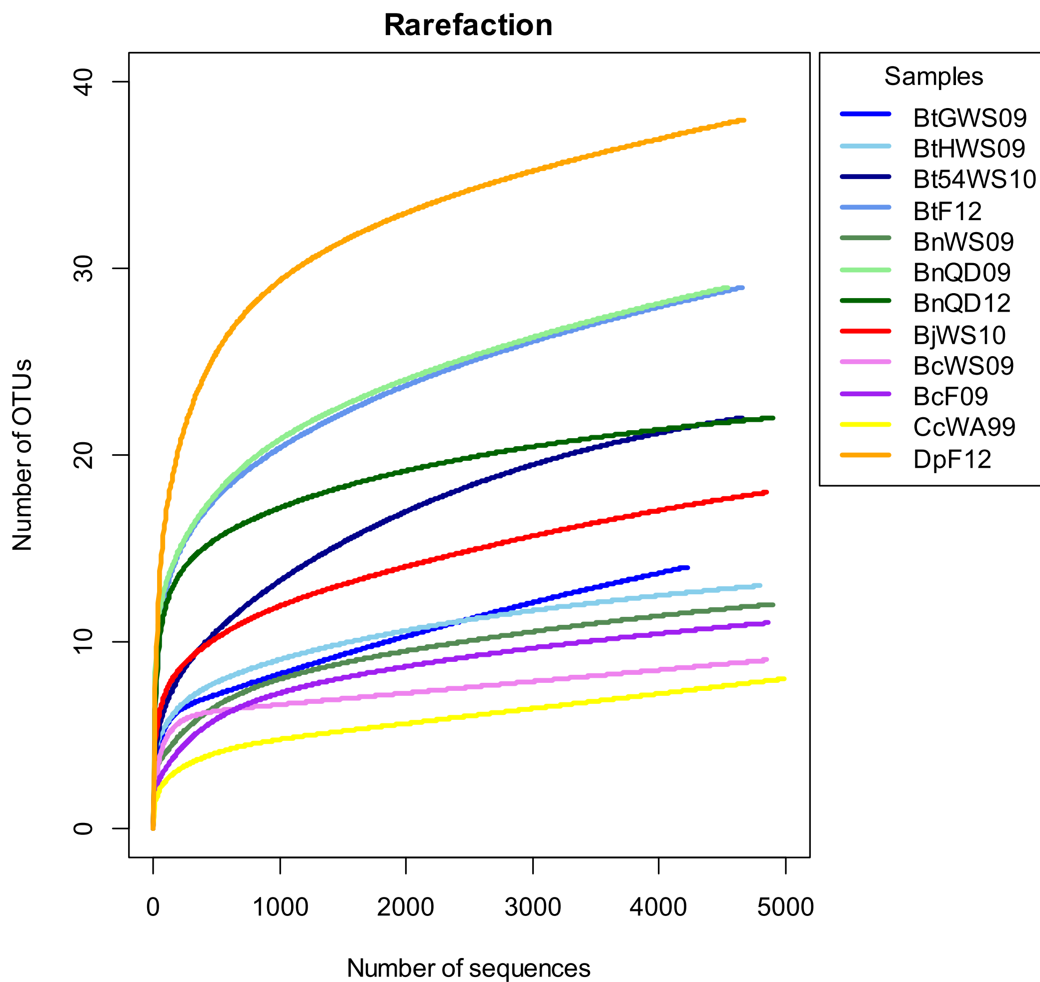



**Figure 4.1** Rarefaction curves of OTUs clustered at 97% similarity for each tephritid fruit fly sample.

*Comparisons of bacterial composition*

Beta diversity measurements were applied in three ways to sequences clustered at 97% similarity. Firstly abundance-based Jaccard distances (Chao *et al.* 2006), applied pairwise to samples, measured abundance of taxa as a fraction of total diversity. Principal coordinates analysis (PCoA) of this metric indicated strong correlation of source (UWS laboratory, QDAFF laboratory and field) and species within the *Bactrocera* group (Figure 4.2). The first two coordinates combined represented 80% of the variation in the data. The microbiome of two sibling species, *B. tryoni* and *B. neohumeralis*, closely related both phylogenetically and by host plant preference, clustered by the source variable more so than by species. The more distantly-related species with more restricted host plant ranges, *B. jarvisi* and *B. cacuminata*, possessed microbiomes more similar to each other, and less similar to *B. tryoni* and *B. neohumeralis* from UWS, despite being maintained in the same environment. Two field-collected samples of *B. tryoni* and *B. cacuminata* from the Sydney region clustered together but separately from their laboratory-reared conspecifics. The two genera *Dirioxa* and *Ceratitis* were most diverged from all of the *Bactrocera* samples (Figure 4.2, Appendix C: Table C.4).

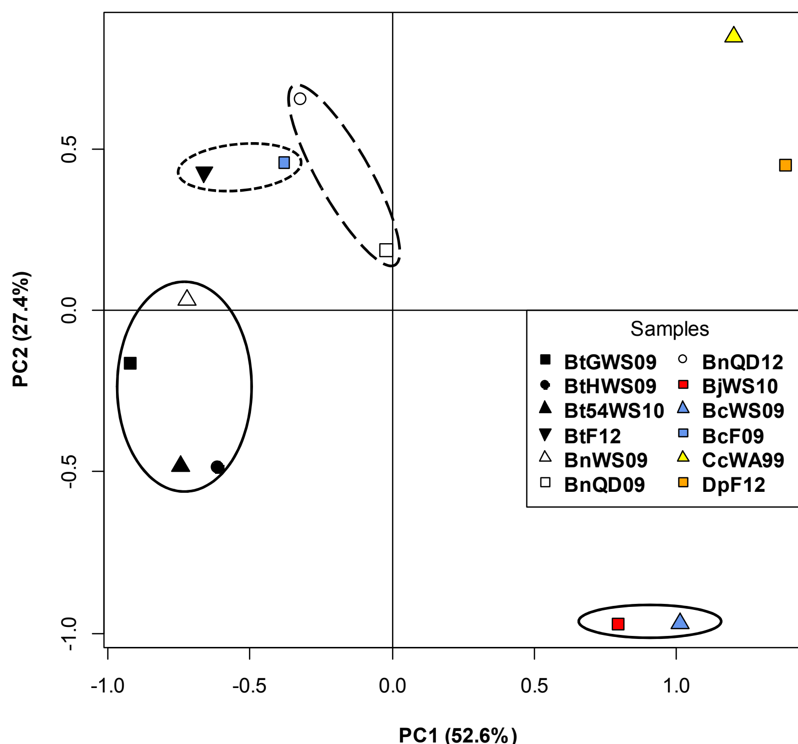



**Figure 4.2** PCoA of Jaccard distances based on OTUs calculated at 97% similarity. Colours represent species. Ellipses highlight similar environments: solid (UWS), long-dashed (QDAFF), short-dashed (field). Axes labels represent the percentage of variation contained in each principal coordinate.

The second beta diversity metric, unweighted UniFrac, measured presence and absence of taxa in a pairwise comparison, however it also accounted for phylogenetic relationships of bacterial taxa. Here, no clear pattern of correlation was produced using PCoA (Figure 4.3, Appendix C: Table C.5) and no UniFrac significance detected ($p > 0.05$). Similarly, the third metric incorporates abundance data for abundance-weighted UniFrac measurement and detected no UniFrac significance ($p > 0.05$), however PCoA was strongly founded on two co-ordinates (PC1 59.4% and PC2 34.1%) and clearly distinguished *D. pornia* from the other species (Figure 4.4, Appendix C: Table C.5). Taken together, these patterns suggested that some of the differences in microbial composition shown by Jaccard distances (notably of BjWS10 and BcWS09 in contrast to the other *Bactrocera* samples) were mitigated by the presence of phylogenetically close but not identical bacterial taxa among the samples examined. The inclusion of abundance data showed that abundance, more so than presence / absence of taxa was responsible for much of the variation observed between species and environments. A heat map was generated for

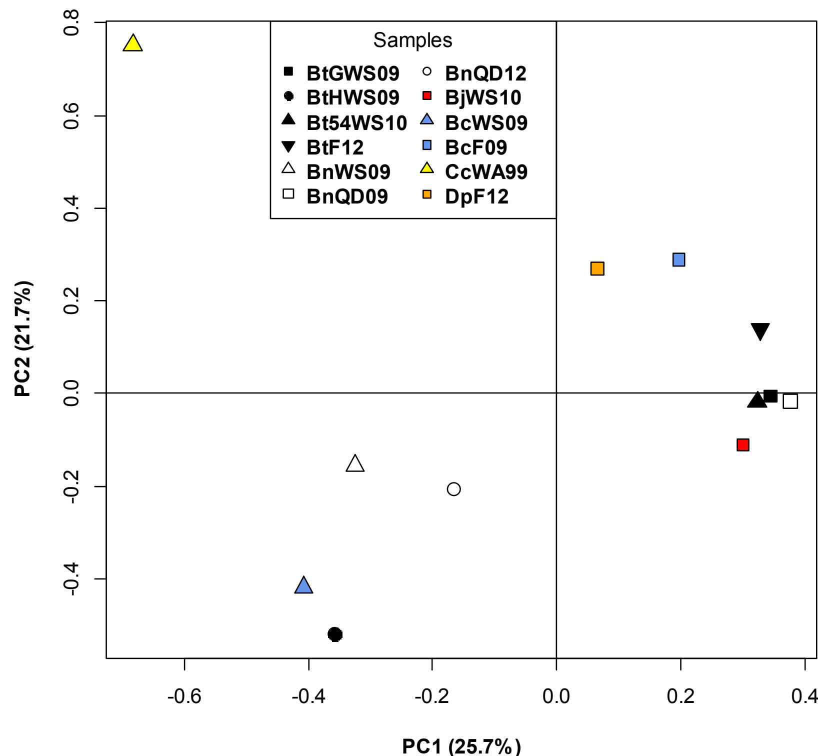



**Figure 4.3** PCoA of unweighted UniFrac distances based on OTUs calculated at 97% similarity. Colours represent species. Axes labels represent the percentage of variation contained in each principal coordinate.

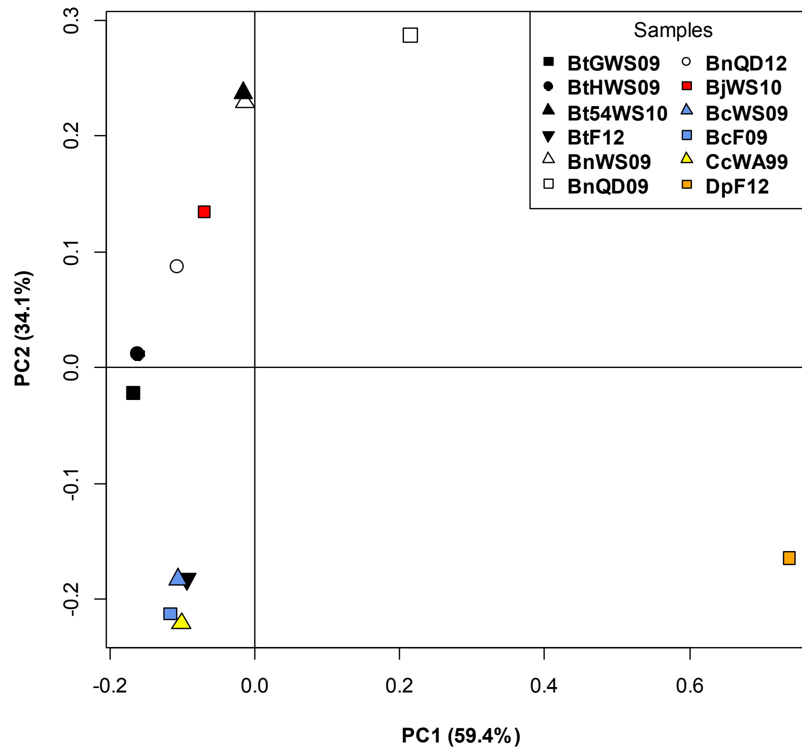



**Figure 4.4** PCoA of weighted UniFrac distances based on OTUs calculated at 97% similarity. Colours represent species. Axes labels represent the percentage of variation contained in each principal coordinate.

the 15 most abundant OTUs at the 95% sequence identity level across the samples (Figure 4.5). The samples were clustered by OTU abundance using Euclidean distances and were compared to the phylogeny of the host species based on the *COI* gene sequences (see Chapter 2). Both *B. tryoni* and *B. neohumeralis* shared mitochondrial haplotype sequences and are ecologically similar, and therefore did not exhibit a detectable species effect on microbial composition.

*Dominant bacteria associated with tephritid fruit flies*

Implementation of RDP Classifier revealed four families that comprised more than 98% of the diversity found overall in these fruit fly species, namely Enterobacteriaceae (68.5%) and Acetobacteraceae (10.2%) of the Proteobacteria phylum; and Enterococcaceae (11.8%) and Streptococcaceae (8.3%) of the Firmicutes phylum (Table 4.2). The remaining sequences represented taxa from five phyla. Individually, all fly samples were dominated (>90%) by one or two of the

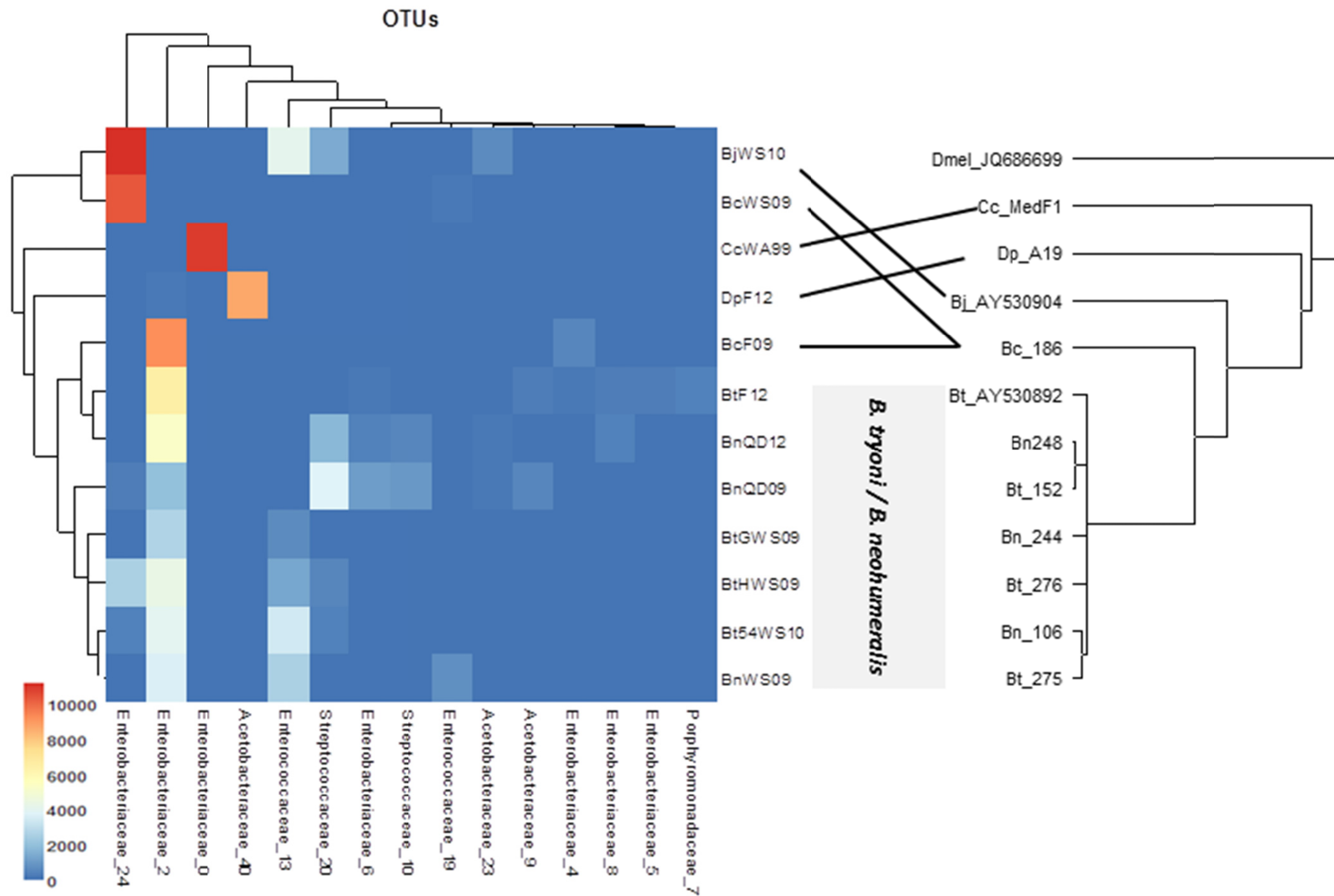



**Figure 4.5** Relationships between hosts and microbial composition. The heatmap shows the relative abundance of the 15 most common OTUs, determined at 95% similarity. The correlation of the dendrogram of bacterial communities in each sample with host *COI* phylogeny is shown. *B. tryoni* and *B. neohumeralis* share *COI* haplotypes and are represented by the grey block. *COI* tree was rooted with *D. melanogaster.*

**Table 4.2** Percentage abundance of family and genus level taxa from each sample, based on RDP Classification. Data are grouped by phyla **(A)** Proteobacteria, **(B)** Firmicutes, **(C)** Bacteroidetes, **(D)** Actinobacteria and **(E)** Chloroflexi.

| **(A) Proteobacteria** | | Sample ID | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Order | **Family** / Genus | BtGWS09 | BtHWS09 | Bt54WS10 | BtF12 | BnWS09 | BnQD09 | BnQD12 | BjWS10 | BcWS09 | BcF09 | CcWA99 | DpF12 |
| Rhodospirillales | **Acetobacteraceae** | **1.34** | **0.37** | **0.61** | **3.75** | **0** | **7.6** | **2.06** | **3.5** | **0.32** | **0.03** | **0** | **92.85** |
| | *Asaia* | 1.34 | | | 3.75 | | 5.35 | | | | | | |
| | *Acetobacter* | | | | | | | | | | | | 87.16 |
| | *Gluconacetobacter* | | | | | | | | | | | | 1.62 |
| | *Gluconobacter* | | | | | | | | | | | | 0.87 |
| | Unclassified | | 0.37 | 0.61 | | | 2.26 | 2.06 | 3.5 | 0.32 | 0.03 | | 3.2 |
| | **Rhodospirillaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.01** | **0** | **0** | **0** | **0** |
| | *Dongia* | | | | | | | | 0.01 | | | | |
| Rhizobiales | **Beijerinckiaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.14** |
| | *Beijerinckia* | | | | | | | | | | | | 0.04 |
| | Unclassified | | | | | | | | | | | | 0.1 |
| | **Brucellaceae** | **0.05** | **0** | **0.12** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Ochrobactrum* | 0.05 | | 0.12 | | | | | | | | | |
| | **Bradyrhizobiaceae** | **0** | **0** | **0.01** | **0** | **0** | **0** | **0** | **0.06** | **0** | **0** | **0** | **0** |
| | *Bosea* | | | 0.01 | | | | | 0.06 | | | | |
| | **Rhizobiaceae** | **0** | **0** | **0.04** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Rhizobium* | | | 0.04 | | | | | | | | | |
| Caulobacterales | **Caulobacteraceae** | **0.03** | **0** | **0** | **0** | **0** | **0.03** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Brevundimonas* | | | | | | 0.03 | | | | | | |
| | *Caulobacter* | 0.03 | | | | | | | | | | | |
| Burkholderiales | **Burkholderiaceae** | **0.03** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Limnobacter* | 0.03 | | | | | | | | | | | |
| | **Comamonadaceae** | **0** | **0** | **0.03** | **0** | **0** | **0** | **0** | **0** | **0** | **0.12** | **0** | **0** |
| | *Delftia* | | | 0.02 | | | | | | | | | |
| | *Pelomonas* | | | 0.01 | | | | | | | | | |
| | *Comamonas* | | | | | | | | | | 0.12 | | |
| | **Oxalobacteraceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.03** |
| | Unclassified | | | | | | | | | | | | 0.03 |
| Aeromonadales | **Aeromonadaceae** | **0.08** | **0** | **0** | **0.01** | **0** | **0** | **0.5** | **0** | **0** | **0** | **0** | **0** |
| | *Aeromonas* | | | | 0.01 | | | 0.5 | | | | | |
| | *Brevundimonas* | 0.08 | | | | | | | | | | | |

| **(A) Proteobacteria** | (cont) | Sample ID | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Order | **Family** / Genus | BtGWS09 | BtHWS09 | Bt54WS10 | BtF12 | BnWS09 | BnQD09 | BnQD12 | BjWS10 | BcWS09 | BcF09 | CcWA99 | DpF12 |
| Enterobacteriales | **Enterobacteriaceae** | **80.31** | **79.03** | **52.28** | **88.5** | **52.22** | **38.64** | **70.42** | **64.09** | **97.31** | **94.77** | **99.99** | **2.42** |
| | *Enterobacter* | 32.4 | 22.09 | 19.94 | 2.35 | 22.46 | 2.33 | 8.11 | | | 2.34 | | 0.1 |
| | *Klebsiella* | 0.03 | 0.02 | 0.05 | | 0.01 | 3.86 | 1.57 | | | | | |
| | *Serratia* | 10.93 | 29.53 | 3.54 | | | 2.95 | | 64.07 | 97.27 | | | 0.03 |
| | *Pantoea* | | | | 0.11 | | | | | | | | 0.03 |
| | *Proteus* | | | | 0.05 | | 0.02 | | | | | | |
| | *Providencia* | | | 0.61 | 1.92 | | 11.79 | 4.86 | | | | | |
| | *Buttiauxella* | | | | 9.71 | | | | | | | | 0.1 |
| | *Kluyvera* | | | | 0.17 | | | | | | | | |
| | *Raoultella* | | | | 1.57 | | 0.15 | 3.54 | | | | | |
| | *Citrobacter* | | | | 10.89 | | | 0.09 | | | | 0.01 | |
| | *Salmonella* | | | | 0.05 | | 0.62 | 7.7 | | | | | |
| | *Morganella* | | | | | | 1.72 | 0.03 | | | | | 0.01 |
| | *Escherichia/Shigella* | | | | 3.9 | | | | | | | | |
| | *Tatumella* | | | | | | | | | | | | 0.04 |
| | Unclassified | 36.95 | 27.39 | 28.14 | 57.79 | 29.74 | 15.19 | 44.5 | 0.02 | 0.04 | 92.41 | 99.98 | 2.12 |
| Legionellales | **Legionellaceae** | **0.03** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Legionella* | 0.03 | | | | | | | | | | | |
| | **Coxiellaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.02** |
| | *Coxiella* | | | | | | | | | | | | 0.02 |
| Pseudomonadales | **Pseudomonadaceae** | **0** | **0** | **0** | **0.25** | **0** | **0.54** | **1.18** | **0** | **0** | **0** | **0.01** | **0** |
| | *Pseudomonas* | 0 | 0 | 0 | 0.01 | | 0.53 | 1.18 | | | | 0.01 | |
| | Unclassified | | | | 0.24 | | 0.01 | | | | | | |
| | **Moraxellaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0.01** | **0** | **0** | **0** | **0** | **0.01** |
| | *Acinetobacter* | | | | | | | 0.01 | | | | | 0.01 |
| Xanthomonadales | **Xanthomonadaceae** | **0.03** | **0** | **0.02** | **0.02** | **0.03** | **0.01** | **0** | **0.07** | **0** | **0** | **0** | **0.77** |
| | *Frateuria* | | | | | | | | | | | | 0.46 |
| | *Stenotrophomonas* | 0.03 | | 0.02 | 0.02 | 0.03 | 0.01 | | 0.07 | | | | |
| | Unclassified | | | | | | | | | | | | 0.31 |
| Gammaproteobacteria_incertae_sedis | | **0** | **0** | **0** | **1.88** | **0** | **0.01** | **0** | **0** | **0** | **4.65** | **0** | **0.18** |
| | *Orbus* | | | | 1.88 | | 0.01 | | | | 4.65 | | 0.18 |

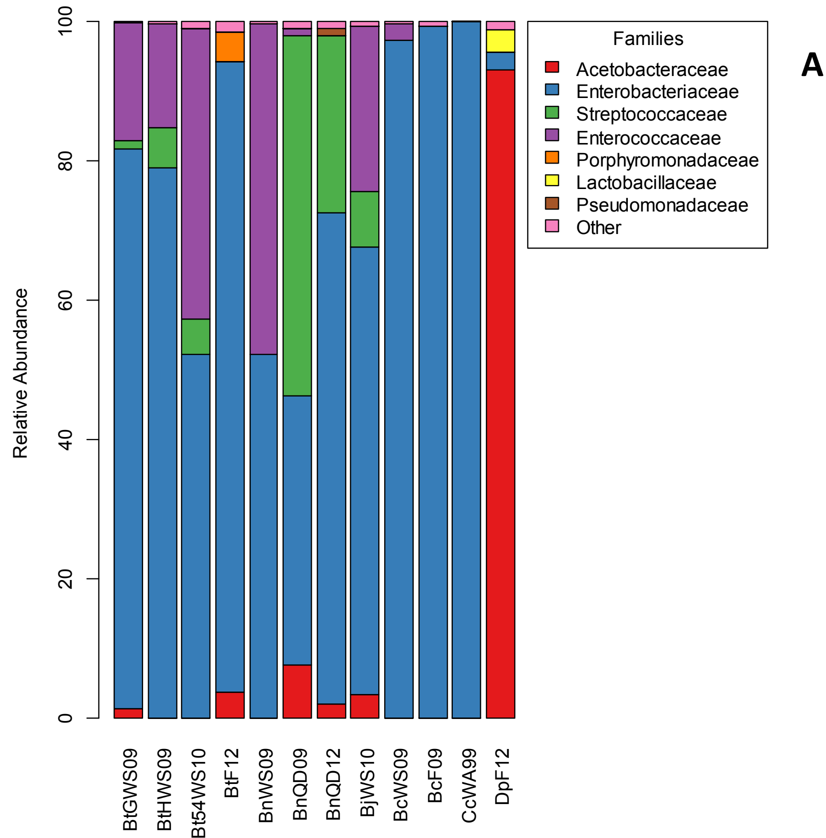
| (B) Firmicutes | | Sample ID | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Order | **Family** / Genus | BtGWS09 | BtHWS09 | Bt54WS10 | BtF12 | BnWS09 | BnQD09 | BnQD12 | BjWS10 | BcWS09 | BcF09 | CcWA99 | DpF12 |
| Lactobacillales | **Enterococcaceae** | **16.95** | **14.92** | **41.64** | **0.02** | **47.37** | **1.06** | **0.41** | **23.85** | **2.37** | **0.22** | **0** | **0.13** |
| | *Enterococcus* | 1.14 | 0.51 | 0.31 | | 10.97 | 0.04 | | 0.5 | 1.7 | 0.22 | | 0.08 |
| | *Vagococcus* | 0.1 | 0.08 | 0.27 | | 0.36 | 0.94 | 0.29 | 0.25 | | | | |
| | Unclassified | 15.71 | 14.33 | 41.05 | 0.02 | 36.04 | 0.08 | 0.11 | 23.09 | 0.67 | | | 0.05 |
| | **Streptococcaceae** | **1.19** | **5.68** | **5.08** | **0.94** | **0.25** | **51.76** | **25.41** | **7.94** | **0** | **0** | **0** | **0** |
| | *Lactococcus* | 1.19 | 5.68 | 5.08 | 0.94 | 0.25 | 51.76 | 25.41 | 7.94 | | | | |
| | **Lactobacillaceae** | **0** | **0** | **0.06** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **3.32** |
| | *Lactobacillus* | | | 0.06 | | | | | | | | | 3.05 |
| | Unclassified | | | | | | | | | | | | 0.27 |
| | **Leuconostocaceae** | **0** | **0** | **0** | **0** | **0** | **0.09** | **0.02** | **0** | **0** | **0** | **0** | **0** |
| | *Fructobacillus* | | | | | | 0.09 | 0.02 | | | | | |
| Clostridiales | **Clostridiaceae 1** | **0** | **0** | **0.02** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Clostridium sensu stricto* | | | 0.02 | | | | | | | | | |
| **(C) Bacteroidetes** | | | | | | | | | | | | | |
| Bacteroidales | **Porphyromonadaceae** | **0** | **0** | **0** | **4.16** | **0** | **0.13** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Dysgonomonas* | 0 | 0 | 0 | 0.69 | | 0.13 | | | | | | |
| | Unclassified | | | | 3.47 | | | | | | | | |
| Flavobacteriales | **Flavobacteriaceae** | **0** | **0** | **0.02** | **0.31** | **0** | **0.14** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Elizabethkingia* | | | 0.02 | | | 0.14 | | | | | | |
| | Unclassified | | | | 0.31 | | | | | | | | |
| Sphingobacteriales | **Sphingobacteriaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.47** | **0** | **0.18** | **0** | **0** |
| | *Sphingobacterium* | | | | | | | | 0.46 | | 0.18 | | |
| | **Chitinophagaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.01** | **0** | **0** | **0** | **0** |
| | *Ferruginibacter* | | | | | | | | 0.01 | | | | |
| **(D) Actinobacteria** | | | | | | | | | | | | | |
| Actinomycetales | **Nocardiaceae** | **0** | **0** | **0.05** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** |
| | *Rhodococcus* | | | 0.05 | | | | | | | | | |
| | **Segniliparaceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.01** |
| | *Segniliparus* | | | | | | | | | | | | 0.01 |
| **(E) Chloroflexi** | | | | | | | | | | | | | |
| Ktedonobacterales | **Ktedonobacteraceae** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0** | **0.02** | **0** | **0** |
| | Unclassified | | | | | | | | | | 0.02 | | |

Unclassified bacteria are assigned to families, but not confidently assigned to genera (>80% bootstrap)

above bacterial families with relative abundance varying among species (Figure 4.6A). When grouping *Bactrocera* samples by environment (UWS, QDAFF and field), only five bacterial families contributed more than 1% of the reads to the composition (Figure 4.6B). Relative abundance of the dominant taxa varied among species; for instance, Enterobacteriaceae were present in all samples, and dominant in all *Bactrocera* samples except BnQD09 (QDAFF) which harboured primarily Streptococcaceae. The most common genera within Enterobacteriaceae (classified to 80% confidence threshold) were *Enterobacter, Serratia, Providencia* and *Klebsiella*. *Enterobacter* were found in all seven *B. tryoni* and *B. neohumeralis* samples; but were not detected in laboratory lines of *B. jarvisi* and *B. cacuminata* where *Serratia* dominated. A large proportion of Enterobacteriaceae in all extracts except *B. jarvisi* and BcWS09 remained unclassified at 80% bootstrap confidence, however at 50% bootstrap confidence most of those bacteria were reclassified as *Enterobacter*. In the family Enterococcaceae, *Vagococcus* and *Enterococcus* were found in seven and nine of the twelve samples, respectively; and *Lactococcus* (Streptococcaceae) was present in eight samples, but highly abundant in the two *B. neohumeralis* samples from QDAFF (Table 4.2).

Other than the four dominant families, an additional 25 families representing 15 orders were detected at low prevalence (Table 4.2); eleven of these families were unique to individual samples. The only Flavobacteriaceae and Pseudomonadaceae were present in three of the seven *B. tryoni* and *B. neohumeralis* samples; *Stenotrophomonas* spp. (Xanthomonadaceae) were detected in three samples of *B. tryoni*, two of *B. neohumeralis* and the *B. jarvisi* sample and in *D. pornia* (as unclassified Xanthomonadaceae).

The bacterial sequences detected in the *C. capitata* sample were almost all Enterobacteriaceae; however none of the sequences could be classified further to genus level with more than 80% bootstrap confidence. When the confidence threshold was relaxed to 50%, sequences were classified as *Enterobacter*, *Citrobacter* and *Erwinia*. The *Enterobacter* sequences in *C. capitata* formed abundant OTUs when clustered at 97% – 100% similarity (98.7% to 57.6% of total reads) so the diversity was low. These OTUs were otherwise found only in field *B. tryoni*, *B. neohumeralis* (from QDAFF) and *D. pornia*, but all at low (<1%)



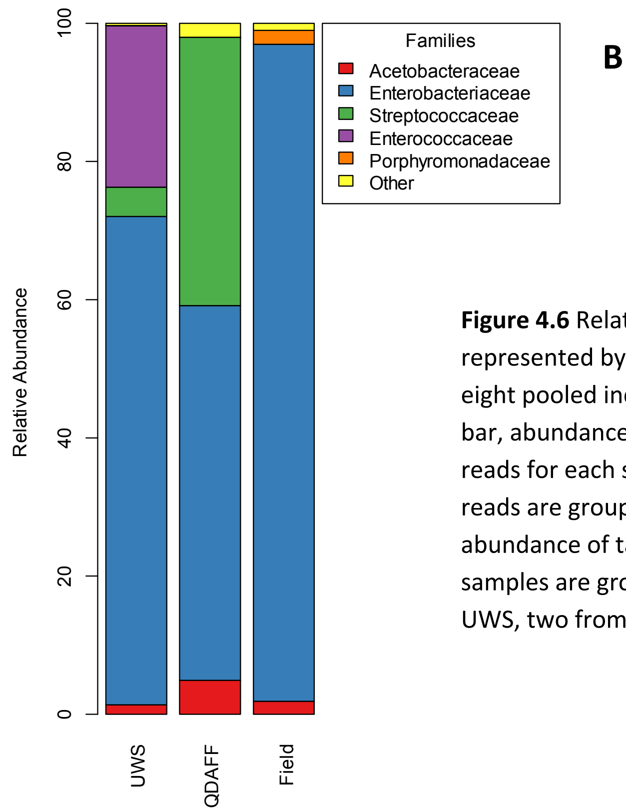




**Figure 4.6** Relative abundance of bacterial families represented by 16S rDNA reads. **(A)** Each sample of eight pooled individuals is represented by a vertical bar, abundance of taxa expressed as a percentage of reads for each sample. Taxa representing <1% of reads are grouped as "other". **(B)** Relative abundance of taxa within the ten *Bactrocera* samples are grouped by environment: six from UWS, two from QDAFF and two field-caught.

prevalence. Unlike the *Bactrocera* samples, no Firmicutes were detected in *C. capitata*, and only a single sequence outside of the Enterobacteriaceae was found, matching *Pseudomonas* (Pseudomonaceae).

Microbial associations of *D. pornia* were strikingly different to *Bactrocera* and *Ceratitis* fruit flies. Over 92% of the *D. pornia* bacterial community was composed of Acetobacteraceae; most belonged to the genus *Acetobacter* (87%), which was absent from all other samples, although other members of the Acetobacteraceae were detected in most *Bactrocera* species, but in relatively low abundance (0-7.6%). This included *Asaia* that was detected at 1-5% in some but not all *B. neohumeralis* and *B. tryoni*.

### 4.5 Discussion

This study is the first to apply 454 pyrosequencing of 16S rDNA amplicons to examine the bacterial communities associated with different tephritid fruit fly species, including significant pest species. It allowed an initial comparison of the relative abundance and composition of bacteria. It also detected rare bacterial taxa within samples belonging to different genera and species with varying geographic distributions, varying degrees of host specialisation, different host plant use, that were sourced from both laboratory and natural environments. Within species we detected streamlining of the microbiome in laboratory-reared flies as opposed to field-collected flies. Across the three tephritid genera, we found that no bacterial genus was common to all three fruit fly genera, or to all *Bactrocera* species, but Enterobacteriaceae were found in all, although at low abundance in *D. pornia*. The latter is a species with different host plant use when compared to the other tephritids included in this study, and, interestingly, was more similar in its microbiome composition to *Drosophila* species that occupy a similar ecological niche. When disregarding the species sampling representation bias in our study, polyphagous *Bactrocera* species had more diverse microbiota than the single represented monophagous *Bactrocera* species. This also confirmed an earlier study that was based on culture-based isolation and physiological characterisation of bacteria (Fitt and O'Brien 1985). Factors such as life stage and adult age impact microbial host

associations (Wang Y *et al.* 2011) ; this was not addressed here but may be a future aspect in understanding the role of microbiomes in *Bactrocera* flies.

*Shared bacterial families in different host species and environments*

This deep sequencing study identified members of the Enterobacteriaceae as the dominant taxa within four Australian *Bactrocera* species. Enterobacteriaceae have also been found consistently in high abundance in these species when using culture and culture-independent techniques (Fitt and O'Brien 1985, Thaochan *et al.* 2010) as well as in *Bactrocera dorsalis* (Wang H *et al.* 2011), *C. capitata* (Aharon *et al.* 2013, Behar *et al.* 2005) and *Rhagoletis* spp. (Howard *et al.* 1985). Firmicutes (primarily of the order Lactobacillales) were also detected in *Bactrocera* species, often at high abundance, using molecular techniques (Thaochan *et al.* 2010, Wang H *et al.* 2011) but rarely in *C. capitata* (Aharon *et al.* 2013). The presence of Enterobacteriaceae and Firmicutes in the microbiome of *Bactrocera* spp., both in laboratory lines and from natural environments, belied the considerable variation observed at the genus and species levels. In our study we observed dominance of *Enterobacter*, *Serratia* and *Citrobacter* in different *Bactrocera* samples, although the limitations of the read length of 454 pyrosequencing hampered assigning genus with certainty in some cases, with an average 36% of Enterobacteriaceae reads unclassified. Fitt and O'Brien (1985) previously found *Serratia*, *Proteus, Klebsiella* and *Enterobacter* as the most abundant genera. Within Firmicutes, Enterococcaceae and Streptococcaceae were prevalent in *B. tryoni*, *B. neohumeralis* and *B. jarvisi*, but in *B. cacuminata* Enterococcaceae were present at low abundance, and Streptococcaceae were absent in both laboratory and field samples. This general phenomenon of differing bacterial composition at genus level between host species, while under the umbrella of consistency at bacterial order and, in some cases, bacterial family levels, has been noted in a study that included microbiome analyses from more than ten *Drosophila* species (Wong *et al.* 2013). Here, we also found no bacterial genera consistently associated with all *Bactrocera* species in all environments. Although the *B. tryoni* and *B. neohumeralis* samples from laboratory and field contained Enterobacteriaceae, the variability in abundance of the bacterial taxa within this

family demonstrated the difficulty of assigning a dominant role to any specific bacterial taxon within this group across all species.

A relationship between *B. oleae* and *Candidatus* E. dacicola is the only recorded obligate bacterial association in *Bactrocera* species thus far (Capuzzo *et al.* 2005), although in some laboratory lines it was not detected through 16S rDNA cloning and sequencing from whole insect genomic extracts (Kounatidis *et al.* 2009), where the dominant bacteria species was *Acetobacter tropicalis*. Very low bacterial diversity was found in laboratory lines of this monophagous fruit fly, which is highly specialised for larval development in the olive fruit. Species richness in the essentially monophagous *B. cacuminata* samples examined in this study was also found to be lower than for the other more generalist *Bactrocera* species, which may suggest that generalist species favour a broader mix of associated bacteria (Prabhakar *et al.* 2013, Wang H *et al.* 2011). We did not detect *Erwinia* and *Acetobacter* in *B. tryoni*, a polyphagous fly capable of also using olives as hosts (Hancock *et al.* 2000). This absence may indicate that the functional roles performed by symbionts in *B. oleae* are either accomplished by other functionally-related bacteria; that the host plant offers the environment to acquire this community; or that the genotype of *B. tryoni* has adapted to the variety of conditions to which the larvae are exposed. It may also indirectly support the notion that *Candidatus* E. dacicola is an obligate symbiont specific to *B. oleae* but not to other *Bactrocera*. Furthermore, a small proportion of the microbiome of *B. tryoni* and *B. neohumeralis* was *Asaia*, an Acetobacteriaceae symbiont commonly found in some mosquitoes (Chouaia *et al.* 2010), and previously also detected in *B. oleae* (Sacchetti *et al.* 2008); while *Acetobacter* may be dominant symbionts of *D. pornia* and *B. oleae* (Kounatidis *et al.* 2009).

*Primary factors affecting the microbiome of tephritid fruit flies*

Variation in microbial composition across biological replicates (three *B. tryoni* lines from UWS; two *B. neohumeralis* lines from QDAFF) was illustrated by the UniFrac metrics (Figure 4.3, 4.4) and similar results were also reported for *Drosophila* samples (Staubach *et al.* 2013). This suggests flexibility in abundance and composition within the physical limits of exposure to bacterial taxa within the

environment. Nevertheless, it appears environment is a primary factor in shaping the bacterial community composition within tephritid hosts. Jaccard analysis presented here showed *B. tryoni* and *B. neohumeralis* UWS samples clustered more closely together than field caught *B. tryoni* from within the Sydney region; and more closely than with two conspecific *B. neohumeralis* lines maintained at QDAFF in Cairns. This largely supports the overriding effect of diet on bacterial composition, which is also consistent with previous *Drosophila* studies (Chandler *et al.* 2011, Cox and Gilmore 2007, Staubach *et al.* 2013).

Within a given environment, whether field or laboratory, the bacterial diversity may also be influenced by host species. The Jaccard distances separated the UWS samples *B. jarvisi* and *B. cacuminata* from *B. tryoni* and *B. neohumeralis*; but UniFrac phylogenetic-based distances showed that the OTUs had some level of sequence similarity. Sibling species *B. tryoni* and *B. neohumeralis* show high genetic similarity in mitochondrial and protein coding nuclear genes (Morrow *et al.* 2000; Gilchrist et al. unpublished), although they are distinguishable through analysis of microsatellite allele frequencies (Wang *et al.* 2003), and this may also account for the similarity of microbiome revealed in samples of these two species from the same source.

The microbiomes of *C. capitata* and *D. pornia* were distinct from the *Bactrocera* species, but the primary driver for their differentiation could not be confidently determined. The *C. capitata* species sits within another tephritid subfamily and is, therefore, phylogenetically separated from *Bactrocera*, but, like *B. tryoni*, is a pest with a large host range. These flies were reared on a different diet in a different environment to all other *Bactrocera* samples; therefore the microbial composition differences may be attributable to species or environment. This *C. capitata* sample was dominated by Enterobacteriaceae, with a single read representing *Pseudomonas*; this low diversity is consistent with a previous deep sequencing study (Aharon *et al.* 2013) and conflicts with the trend of reduction in microbial diversity seen from polyphagous to monophagous *Bactrocera*. Furthermore, the dissimilarities of *Ceratitis* and *Bactrocera* were displayed by the differences of OTU abundance and composition at higher taxonomic levels.

The *D. pornia* microbiome was dominated by the bacterial family Acetobacteraceae, in particular *Acetobacter*, and was distinct from all other samples except when

unweighted UniFrac was applied, and then it was more similar to *B. tryoni* and *B. cacuminata* field samples. This finding indicates that microbial composition is taxonomically related among the different genera analysed in this study, but differences in abundance are detected. Although *D. pornia* has a large host fruit range (Hancock *et al.* 2000), all reports are for infestation of fallen or damaged fruit which sets this species apart from the pest species such as *B. tryoni* that oviposit into ripening fruit. Acetobacteraceae were also an abundant component of the bacterial community in *Drosophila* species (Cox and Gilmore 2007, Staubach *et al.* 2013, Wong *et al.* 2011), which share a similar, fermentation-driven ecological niche with *D. pornia*. However, in *Drosophila*, variation in presence and abundance of particular bacteria, such as *Acetobacter*, was extreme, not only across different environments and fly species within an ecological niche, but also across replicate samples (Wong *et al.* 2013).

Overall, the dominance of Enterobacteriaceae, Acetobacteriaceae and two Lactobacillales families, Enterococcaceae and Streptococaceae, in *Bactrocera*, *Ceratitis* and *Dirioxa*, aligned neatly with the common microbial composition of *Drosophila* species (Chandler *et al.* 2011, Wong *et al.* 2011, Wong *et al.* 2013), with the exception that Lactobacillaceae are more common than Streptococcaceae in *Drosophila*. Differences in abundance, however, were substantial, and may be attributed to three factors. Environmental access to particular bacterial taxa was a significant factor, resulting from larval ingestion, following bacterial inoculation and amplification in the fruit during and after oviposition (Behar *et al.* 2008a). Usage of prior oviposition sites was observed in *B. tryoni* and may increase the diversity of bacteria accessible to developing larvae (Christenson and Foote 1960). Host physiology also exerted some control, accounting for the species-specific and ecological-niche effects primarily on abundance of taxa. Opportunistic colonisation may result from accessibility to particular taxa as the cuticular gut lining is shed at pupation (Murphy *et al.* 1994); or opportunistic increases in pathogenic taxa following the decrease in abundance of a dominant taxon (Capuzzo *et al.* 2005, Ryu *et al.* 2008) may account for the extensive variability in microbial composition and abundance, particularly noted among individuals of the same laboratory line (Staubach *et al.* 2013).

*Absence of Wolbachia*

*Wolbachia* are common endosymbionts, found in approximately 40% of insect species and are primarily, but not exclusively, located in the gonads. *Wolbachia* is widespread in *Drosophila* populations, and some studies have detected *Wolbachia* in such high abundance using cloning and 454 pyrosequencing methods to the extent that infected host taxa were eliminated from some analyses (Cox and Gilmore 2007, Staubach *et al.* 2013). Other studies either included *Wolbachia* sequences (Corby-Harris *et al.* 2007) or filtered out *Wolbachia* sequences (Wong *et al.* 2013). When *Wolbachia* sequences were removed prior to analysis, comparative analyses with and without *Wolbachia* showed similar microbial diversity (Wong *et al.* 2013). neither does the presence of *Wolbachia* in some populations or laboratory cultures under the same conditions alter the remaining microbial diversity (Staubach *et al.* 2013). *Wolbachia* was not detected in laboratory lines using standard and high-sensitivity PCR-based methods (Chapter 2) or by using deep sequencing protocols. Specimens used in this study from field-collected samples from temperate climates have also been screened using the same PCR technique and 454 pyrosequencing without any detection of *Wolbachia* sequences. However, multiple *Wolbachia* infections were detected at low prevalence in tropically-located *B. neohumeralis* and *B. tryoni* field-caught individuals (Chapter 2). It will be useful to further assess the microbiome of such infected flies from tropical regions in order to compare the richness and diversity.

The detection of rare taxa is an advantageous feature of deep sequencing protocols, and it confirmed the absence of *Wolbachia* from any of the lines or populations examined here. Although Aharon *et al*. (2013) did not report *Wolbachia* for the microbiome of *C. capitata* samples collected in Israel, we found *Wolbachia* sequences at low frequency in their raw dataset deposited at MGRAST (http://metagenomics.anl.gov/linkin.cgi?project_2348); however, previous PCR-based screening methods had failed to detect *Wolbachia* in any natural populations of *C. capitata* (Zabalou *et al.* 2004) except from Brazil (Rocha *et al.* 2005).

*Applications to pest control*

Detailed knowledge of the microbial associations of pest insects can assist in the development of control strategies that are more effective, species-specific and environmentally responsible. Assessment of the microbes required by fruit fly species for optimal fitness can be exploited; for instance attraction to certain bacteria as food sources or as oviposition sites (Drew *et al.* 1983, Drew and Lloyd 1987, Lauzon *et al.* 2000, Meats *et al.* 2009, Prabhakar *et al.* 2013, Thaochan and Chinajariyawong 2011, Wang *et al.* 2013) may be utilised in the optimisation of lures and baits for both male and female annihilation (Clarke *et al.* 2011, Dominiak and Ekman 2013). Provisioning of beneficial microbes in the form of probiotic diets to larvae could reduce costs for fruit fly rearing, while probiotic diets for adults could improve the fitness of irradiated sterile SIT or *Wolbachia* infected IIT flies that may have lost microbial diversity during generations of laboratory adaptation or following radiation treatment (Ben Ami *et al.* 2010, Sacchetti *et al.* 2014). Alternatively, denying fruit flies favourable microbes that provide nutritional benefit or even insecticide breakdown (Bousch and Matsumura 1967) may be employed to improve efficacy of current control methods.

Paratransgenesis – the genetic modification of symbiotic bacteria to induce expression of particular phenotypes within their hosts – is another promising avenue of research for pest control. Application to arthropod vectoring of disease has been explored (Beard *et al.* 1993, Weiss and Aksoy 2011) but paratransgenesis may also be applied to pest control through delivery of antimicrobial products that destroy the natural assemblage of bacteria within the gut. This is particularly useful for the fitness of insects that rely on their gut microbiome for nutrient acquisition, such as termites (Husseneder and Collier 2009). Alternatively, such bacteria may be manipulated to deliver insecticidal dsRNA (or sex-ratio manipulating dsRNA e.g. of sex-determination genes) by ingestion in a highly species-specific manner (Tian *et al.* 2009, Whyard *et al.* 2009). The potential of orally delivered RNAi has been investigated through identification of important transcripts involved in the RNAi mechanism expressed in the midgut of tephritid fruit fly *Anastrepha obliqua* (Shen *et al.* 2013). However, some applications require embryonic targets, and would depend upon reliable maternal transmission to early embryos. Vertically transmitted *Wolbachia* may be useful here, but no successful transformation of *Wolbachia* has

been achieved by any researchers to date. In *B. oleae*, bacteria are deposited on the egg surface during passage through the ovipositor, and enter the embryo through the micropyle (Capuzzo *et al.* 2005, Petri 1910), thus providing an early embryonic delivery system for this species. This has potential for bacterial-mediation of RNAi-induced female-lethality or female-to-male conversion in the production of male-only fruit fly lines, crucial to the efficacy of IIT release cohorts and of relevance to efficient SIT.

*Conclusions*

The objectives of this study were to uncover the bacterial communities associated with tephritid fruit flies in Australia and analyse the primary drivers of taxonomic composition. Four bacterial families were consistently found in tephritid flies, very similar to *Drosophila* at this taxonomic level, but differences within these bacterial families and also in their abundance were substantial. Low abundance taxa were found, but none were consistently present in any species, environment or ecological group. Species comparisons within the genus *Bactrocera*, across genera and across ecological niches revealed that the particular environment, probably also typified by the diet, seemed to be an important determinant of microbial composition in closely-related species; beyond this, host species identity or ecology also had an effect.

Studies so far have failed to find a consistent microbiome, at bacterial genus or species level, associated with the *Bactrocera* fruit flies examined here, which may suggest flexibility in the bacterial community associated with these flies. This may be useful for the development of pest management strategies that rely on bacterial manipulation. However, previous studies on bacteria as a food source and attractant for *Bactrocera* species delivered contradictory findings (Drew *et al.* 1983, Meats *et al.* 2009), which may imply that very particular strains of bacteria producing particular metabolites, are required as effective lures and probiotics. Some of this may also reflect the limitations of bacterial taxonomy based on single locus approaches, including 16S rDNA analysis, (which frequently suffers from a lack of true biological and technical replication, see Discussion 7.2.3) or high metabolic diversity within some bacterial taxa. Metagenomic and metabolomics approaches rather than amplicon pyrosequencing will be avenues for future research to overcome

this issue. Furthermore, better understanding of the transmission of bacteria from mother to offspring is necessary, as culture-dependent methods found less diversity in eggs from four *Bactrocera* species (Fitt and O'Brien 1985), and in *C. capitata* not all bacteria were found in all life stages, notably the absence in eggs of *Enterobacter* spp., which are highly abundant in adults (Behar *et al.* 2008a). Delivery of bacteria for different targets will require a comprehensive picture of optimal bacterial composition at different life stages of each species of interest.

# Chapter 5

# Expression patterns of sex-determination genes in single male and female embryos of two *Bactrocera* fruit fly species during early development

Morrow, J. L., Riegler, M., Frommer, M., and Shearman, D. C. A. “Expression patterns of sex-determination genes in single male and female embryos of two *Bactrocera* fruit fly species during early development” Insect Molecular Biology (in preparation).

## 5.1 Abstract

In tephritid fruit flies, the sex-determination pathway follows the auto-regulated sex-specific splicing of *transformer* mRNA, and the cooperation of *transformer* (*tra*) and *transformer-2* (*tra*-2) to effect the sex-specific splicing of *doublesex* (*dsx*), the genetic double-switch responsible for male or female somatic development. In control of this is the *Dominant Male Determiner* (*M*) that functions as the primary sex-determination signal. *M* is as yet uncharacterised, but is Y-chromosome linked, expressed in the zygote and either directly or indirectly reduces the amount of active TRA protein in male embryos. Here we first demonstrated the conservation of gene and transcript structure of *tra*, *tra-2* and *Sex-lethal* (*Sxl*) explicitly in two Australian tephritid fruit fly species, *Bactrocera tryoni* and *Bactrocera jarvisi*. We then used a single, sexed embryo approach to examine, by quantitative RT-PCR, the time course of expression of important known genes in the sex-determination pathway in individual male and female *Bactrocera* during early embryonic development. Embryo sexing of a *B. jarvisi* line, plus a *B. tryoni* line with an introgressed *B. jarvisi* Y-chromosome, was possible due to molecular markers located on the *B. jarvisi* Y-chromosome. Results showed that, in *B. jarvisi*, *M* actively influenced expression of sex-specific *tra* transcripts between 3 to 6 hours after egg laying, and the *dsx* isoform was established by 7 hours. These milestones were delayed in the *B. tryoni* hybrid line with the *B. jarvisi* Y-chromosome. The results provide a developmental time frame for transcriptomic analyses, plus information on genes that may be targeted for the development of male-only strains useful in pest management strategies.

### 5.2 Introduction

Studies of sex determination in insects explore the fascinating developmental regulation required to convert a haploid egg containing maternal products to a fertilised embryo that must transition from maternal to sex-specific zygotic control. The differences in the primary sex-determination signals among the insect taxa at family, genus and even species level (Saccone *et al.* 2002, Schutt and Nothiger 2000, Shearman 2002) provide a means to examine the evolution of this fundamental pathway. For many pest species, sex-determination genes are also valuable targets for the development of genetic-sexing lines, which have hitherto successfully improved pest-management strategies for *Ceratitis capitata* (Diptera: Tephritidae), notably the sterile insect technique (SIT; Hendrichs *et al.* 1995) and incompatible insect technique (IIT; Zabalou *et al.* 2009).

Sex determination in *Drosophila melanogaster* Meigen (Diptera: Drosophilidae) is well studied, and follows a pathway completed by differential splicing of the *doublesex* (*dsx*) mRNA, resulting in production of male and female protein isoforms (reviewed in Clough and Oliver 2012). The first tephritid in which male and female-specific splicing of *dsx* was shown was Australian pest species, *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae) (Shearman and Frommer 1998). Like *D. melanogaster*, the carboxyl termini of *Bt*DSX differ between adult males and females, with a common region at the amino terminus; similarly the mRNA was found to be alternatively spliced in males and females, with *dsx* repeat elements (*dsx*RE) located in a non-coding region of the female-specific exon. As demonstrated in *D. melanogaster*, the female splicing mode of *dsx* transcripts is facilitated by *transformer* (TRA) and *transformer-2* (TRA-2) proteins binding to *dsx*RE; without TRA and TRA-2 the male transcript is produced (Tian and Maniatis 1993).

Transcripts of *tra* are also spliced into male and female forms, and only the female transcripts encode a functional protein. In *C. capitata* (Pane *et al.* 2002), *Bactrocera oleae* (Lagos *et al.* 2007) and *Anastrepha* species (Ruiz *et al.* 2007), male and female-specific splicing of *tra* homologues was demonstrated. In these species the pre-mRNA transcripts of *tra* contain TRA/TRA-2 binding sites like in *dsx*, but here they are found in a region covering male-specific exons and their interspersed introns. In contrast to *dsx*, the binding of TRA and TRA-2 to the *tra* pre-mRNA acts

to block strong canonical splice sites, leading to the use of the weak female-specific splice sites and the removal of the strong splice sites and associated exons to yield the female-specific mRNA. Translation of the female-specific transcript encodes a longer protein product, allowing an auto-regulated supply of female-specific *tra* mRNA and $TRA^F$ protein to be maintained in cells. Male-specific *tra* transcripts incorporate stop codons early, leading to production of a truncated and theoretically non-functional protein.

In tephritids, TRA-2 is also essential, as part of the spliceosome complex with TRA, for correct splicing of *tra* and *dsx* transcripts, however it is expressed in both males and females during embryogenesis. Elimination of either *tra* or *tra-2* by RNAi early in development disrupts the mechanism favouring female-specific splicing of *tra* and *dsx* in *C. capitata* and *Anastrepha* spp. (Pane *et al.* 2002, Salvemini *et al.* 2009, Sarno *et al.* 2010, Schetelig *et al.* 2012) and leads to the expression of $dsx^M$ and the consequent development of male somatic tissue (Saccone *et al.* 2008).

In drosophilids, the *Sex-lethal* (*Sxl*) gene constitutes the apex of the sex-determination pathway, being transcribed and processed into sex-specific forms, in response to the number of X-chromosomes (Erickson and Quintero 2007). However, *Sxl* does not produce different transcripts in males and females in tephritids (Lagos *et al.* 2005, Saccone *et al.* 1998) and other Diptera, such as *Musca domestica* (Meise *et al.* 1998) and *Megaselia scalaris* (Sievert *et al.* 2000), but is transcribed during early development in *C. capitata* (Gabrieli *et al.* 2010) and is highly expressed in the pole cells of the syncytial blastoderm (Saccone *et al.* 1998). Furthermore, the presence of the Y-chromosome and not the number of X-chromosomes determines the gender of *C. capitata* flies (Willhoeft and Franz 1996a, Willhoeft and Franz 1996b). The putative *Dominant Male Determiner* (*M*), which is the apex of the sex-determination pathway in tephritids, has been localised to the long arm of the *C. capitata* Y-chromosome (Willhoeft and Franz 1996b), but has yet to be characterised in any insect species that rely on this sex-determination system.

Here we used a single-embryo approach to analyse the time course of expression of some important genes in the sex-determination pathway during early embryonic development. Distinguishing male and female embryos was critical to the analysis. Hence, we utilised *Bactrocera jarvisi* (Tryon) (Diptera: Tephritidae), a widespread

Australian fruit fly and emerging pest (Fay 2012, Hancock *et al.* 2000), which possesses a known Y-chromosome genetic marker and an ability to produce fertile hybrids with *B. tryoni* under laboratory conditions (Shearman *et al.* 2010). We also constructed an introgressed *B. tryoni*:*B. jarvisi* hybrid line, which retained the *B. jarvisi* Y-chromosome in an otherwise complete *B. tryoni* genetic background. Our objectives were to examine and compare sex determination in two *Bactrocera* species, determine any differences between male and female expression of common transcripts; differences in expression of sex-specific transcripts; and to use these data to estimate at what time during early embryogenesis and infer in what manner *M* acts to influence sex-specific splicing, leading to male and female differentiation.

### 5.3 Results

*Sex-determination genes in* Bactrocera *spp.*

Gene homologues of the sex-determination pathway as described in *D. melanogaster* and *C. capitata* were targeted for this study. DNA sequence for both male and female transcripts of *dsx* was available (Shearman and Frommer 1998), but characterisation was required for *tra*, *tra-2* and *Sxl* in *B. tryoni* and *B. jarvisi*.

For *tra* and *tra-2* genes, genomic and mRNA sequences were acquired from adult male and female *B. tryoni* and *B. jarvisi*. The gene structures and sequences of *Bttra* and *Bjtra* are very similar to each other; and contain five common exons that correspond to exons 1a, 1b, 2a, 2b and 3 in *B. oleae* (Lagos *et al.* 2007); and male-specific exons MS1 and MS2 located between exon 1b and exon 2a (Figure 5.1). Transcripts of *Bttra* and *Bjtra*, like *tra* from *C. capitata* (Pane *et al.* 2002) and *B. oleae* (Lagos *et al.* 2007), were alternatively spliced in males and females, with females of *B. tryoni* generating a transcript encoding 408 amino acids and *B. jarvisi* 422 amino acids (Table 5.1). Male transcripts were larger, due to incorporation of MS1 and MS2 sequences, through which in-frame stop codons were introduced. This shorter male ORF encodes a 66 amino acid protein devoid of the important serine-arginine rich region found in the female isoform of *Bt*TRA, *Bj*TRA and in previously studied tephritids and drosophilids (Boggs *et al.* 1987, McKeown *et al.* 1987, Pane *et al.* 2002). Putative TRA/TRA-2 binding sites were identified in the male-specific

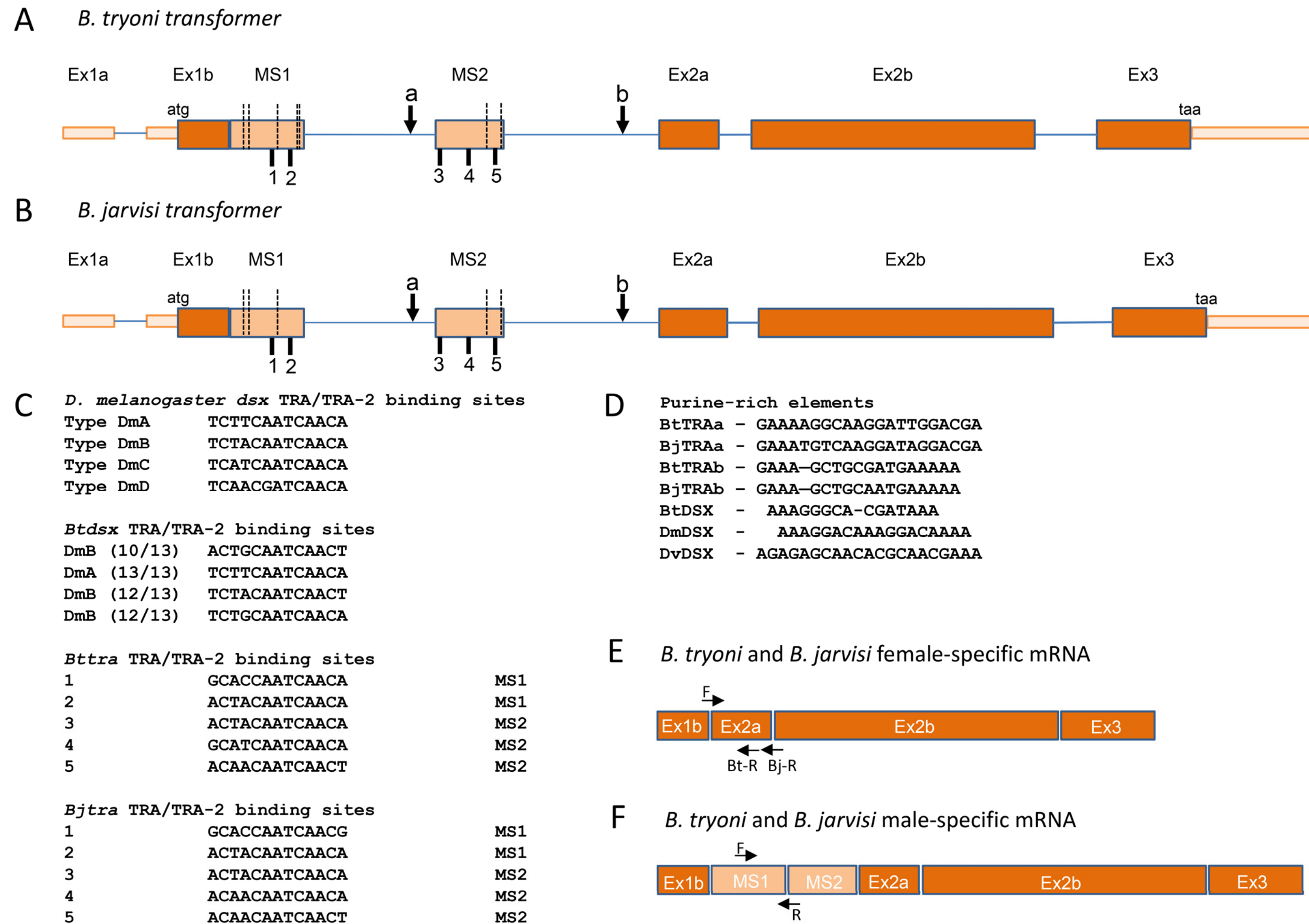



**Figure 5.1** See next page for legend

**Figure 5.1** (previous page) Genetic structure of the transformer gene of *B. tryoni* **(A)** and *B. jarvisi* **(B)**. Dark orange exons are common to male and female spliced transcripts, light orange are male-specific. Thin rectangles are untranslated regions. Locations of putative TRA/TRA-2 binding sites are numbered and putative purine-rich elements are lettered. Stop codons introduced into male transcripts in MS1 and MS2 are represented by dashed lines. **(C)** TRA/TRA-2 canonical sequences from *D. melanogaster* and putative sequences from *B. tryoni dsx* (Shearman and Frommer, 1998) and *Bttra* and *Bjtra*. **(D)** Purine-rich elements in *D. melanogaster* and *D. virilis dsx* genes, and similar sequences in *Btdsx* (Shearman and Frommer, 1998), *Bttra* and *Bjtra*. **(E)** Representation of the female-specific mRNA transcript of *B. tryoni* and *B. jarvisi*; arrows define the location of the qRT-PCR primers (forward and reverse; note: common forward primer only). **(F)** Representation of the male-specific mRNA transcript of *B. tryoni* and *B. jarvisi*; arrows define the location of the qRT-PCR primers (forward and reverse).

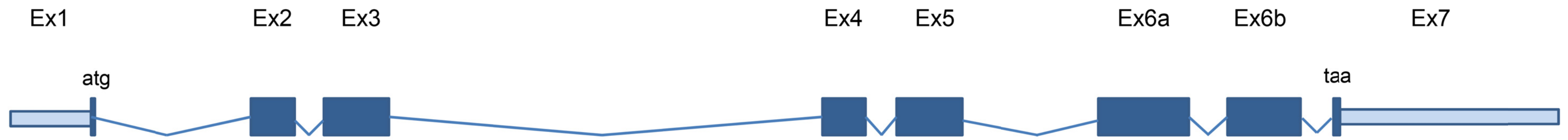
**Table 5.1.** *tra* female-specific amino acid percentage identity between species (below diagonal), expressed as a fraction of total non-gap sites only (above diagonal). Sequences of other species were obtained from NCBI: *B. oleae* (AJ15413), *C. capitata* (AF434936), *A. obliqua* (EU024498) and *D. melanogaster* (NM079390). Number of amino acids refers to the complete ORF of the female-specific (or non-sex-specific in *D. melanogaster*) TRA protein.

| Species | *B. tryoni* | *B. jarvisi* | *B. oleae* | *C. capitata* | *A. obliqua* | *D. melanogaster* |
|---|---|---|---|---|---|---|
| *B. tryoni* | | 28/408 | 73/407 | 167/402 | 168/390 | 124/172 |
| *B. jarvisi* | 93.1 | | 85/421 | 174/416 | 178/404 | 126/172 |
| *B. oleae* | 82.1 | 79.8 | | 171/417 | 172/405 | 121/172 |
| *C. capitata* | 58.5 | 58.2 | 59.0 | | 179/411 | 126/173 |
| *A. obliqua* | 56.9 | 55.9 | 57.5 | 56.4 | | 129/174 |
| *D. melanogaster* | 27.9 | 26.7 | 29.7 | 27.2 | 25.9 | |
| No. amino acids | 408 | 422 | 422 | 429 | 417 | 197 |

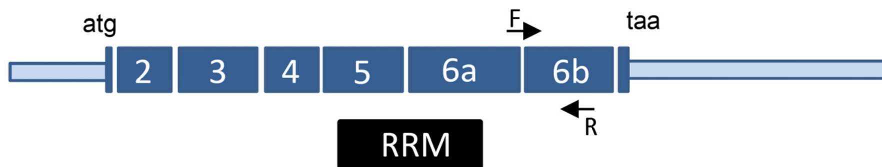
exons and two putative purine-rich elements (PRE) were identified in the introns of both *Bttra* and *Bjtra* (Figure 5.1).

The *tra-2* gene comprises eight exons, is highly conserved between *B. tryoni* and *B. jarvisi* and is structurally comparable to *B. oleae* (GenBank Accession Number AJ547623), *C. capitata* (Salvemini *et al.* 2009), *M. domestica* (Burghardt *et al.* 2005) and *D. melanogaster* (Amrein *et al.* 1988, Mattox *et al.* 1990) (Figure 5.2). *Bttra-2* and *Bjtra-2* transcripts were common to male and female flies, and produced an ORF encoding 255 and 251 amino acids respectively (99.6% amino acid identity; Table 5.2). The most highly conserved region is the RNA recognition motif (RRM), a stretch of 72 amino acids identical in the *Bactrocera* species, different by a single amino acid to *C. capitata*, and with 69.4% (50 of 72aa) identity to *D. melanogaster* (Figure 5.2).

A *B. tryoni* and *B. jarvisi transformer-2* genomic structure




B *B. tryoni* and *B. jarvisi* common mRNA transcript




C *B. jarvisi* Y-chromosome genomic structure

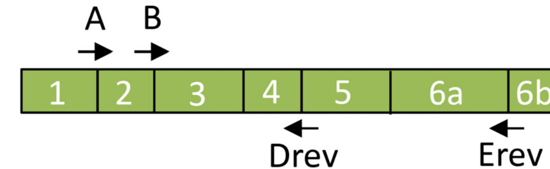



**Figure 5.2 (A)** The genetic structure of the *tra-2* gene in *B. tryoni* and *B. jarvisi* (approximately to scale). Dark blue boxes are exons, thin rectangles are untranslated regions, lines represent introns. **(B)** Representation of the mRNA transcript. Location of conserved RRM (RNA recognition motif) in protein is shown; arrows define the location of the qRT-PCR primers (forward and reverse). **(C)** Representation of the *B. jarvisi* Y-chromosome *tra-2* pseudogene genomic DNA structure. Arrows identify forward and reverse primers designed to amplify this DNA specifically, apart from the autosomal *tra-2* gene.

**Table 5.2** *tra-2* amino acid percentage identity (below diagonal) and fraction of differences (above diagonal) between species. Other *tra-2* sequences were retrieved from NCBI: *B. oleae* (DQ100253), *C. capitata* (EU999754) and *D. melanogaster* (NM057416). Number of amino acids refers to the length of the complete ORF of the TRA-2 protein in each species.

| Species | *B. tryoni* | *B. jarvisi* | *B. oleae* | *C. capitata* | *D. melanogaster* |
|---|---|---|---|---|---|
| *B. tryoni* | | 1/251 | 9/251 | 23/250 | 132/244 |
| *B. jarvisi* | 99.6 | | 9/251 | 24/250 | 132/244 |
| *B. oleae* | 96.4 | 96.4 | | 27/250 | 132/244 |
| *C. capitata* | 90.8 | 90.4 | 89.2 | | 136/244 |
| *D. melanogaster* | 45.9 | 45.9 | 45.9 | 44.3 | |
| No. amino acids | 255 | 251 | 251 | 251 | 264 |

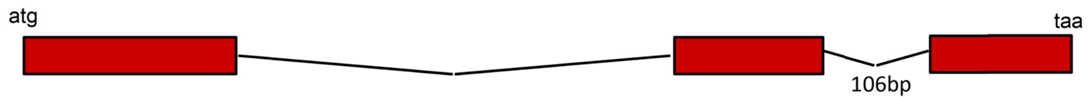
Partial mRNA sequence for *Sxl* was acquired by amplifying male and female adult *B. tryoni* and *B. jarvisi* cDNA using primers designed from the *B. oleae Sxl* (GenBank Accession No. AJ715415). The same transcripts were present in both sexes of each species, and encoded a putative polypeptide of 339 amino acids which displayed particularly high homology to other tephritid *Sxl* and *D. melanogaster Sxl* in the two RRM regions (Table 5.3, Figure 5.3).

**Table 5.3** *Sxl* amino acid percentage identity (below diagonal) and fraction of differences (above diagonal) between species. Other *Sxl* sequences were obtained from NCBI: *B. oleae* (AJ715415), *C. capitata* (AF026145) and *D. melanogaster* (M23636). Number of amino acids refers to the length of the ORF of the longest SXL protein in each species.

| Species | *B. tryoni* | *B. jarvisi* | *B. oleae* | *C. capitata* | *D. melanogaster* |
|---|---|---|---|---|---|
| *B. tryoni* | | 6/302 | 10/333 | 17/329 | 77/324 |
| *B. jarvisi* | 99.3 | | 8/301 | 14/297 | 70/293 |
| *B. oleae* | 97.9 | 98.3 | | 11/333 | 72/324 |
| *C. capitata* | 96.4 | 96.6 | 97.9 | | 73/322 |
| *D. melanogaster* | 84.9 | 83.4 | 85.2 | 84.8 | |
| No. amino acids | 339 | 307* | 339 | 340 | 354 |

*The full *B. jarvisi* ORF was not obtained, sequence begins at *B. tryoni* amino acid no. 37

A Partial *B. tryoni* and *B. jarvisi Sex-lethal* genomic structure




B Partial *B. tryoni* and *B. jarvisi* mRNA transcripts

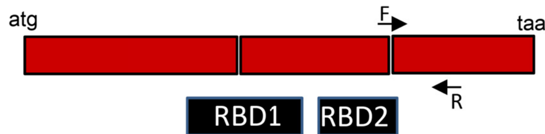



**Figure 5.3 (A)** Partial genetic structure of the *Sxl* gene in *B. tryoni* and *B. jarvisi*, exons are red rectangles, introns are lines. **(B)** The complete coding sequence in *B. tryoni* has been sequenced. The locations of two conserved RNA binding domains (RBD1 from amino acid 108-186 and RBD2 from amino acid 194-246) are displayed. Arrows define the location of the qRT-PCR primers (forward and reverse) used in both *B.tryoni* and *B. jarvisi*. Exons are not numbered as we lack the entire genomic structure.

A target gene that is not expressed maternally, but is expressed zygotically during early embryogenesis was required to signal zygotic transcription and confirm that development had proceeded. The gene, *slow as molasses* (*slam*), involved in cellularisation and expressed in this way in both *Drosophila* and *C. capitata* (Gabrieli *et al.* 2010), was chosen. The sequence of *slam* was obtained by amplification of cDNA from male and female *B. tryoni* using primers designed from the *B. tryoni* genome (Gilchrist *et al*. unpublished). The transcript sequences derived from this region were common in both males and females.

*Y-chromosome markers in* B. jarvisi

*Bactrocera jarvisi* males are known to carry mitochondrial DNA markers on the Y-chromosome (Shearman *et al.* 2010). These allow for discernment of the sex of embryos otherwise lacking morphologically distinguishing characters. Unexpectedly, PCR amplification of male *B. jarvisi* genomic DNA with *tra-2* primers generated smaller *tra-2* fragments than those amplified from adult females and from both sexes in *B. tryoni*. Sequencing of the genomic PCR products revealed an intronless *tra-2* pseudogene from exon1 to exon 6b, likely situated on the Y-chromosome as it was

only amplified from males (Figure 5.2C). Primers were designed to specifically amplify this sequence by locating the forward and reverse primers on the exon junction sites in the autosomal gene. These primers (Appendix D: Table D.1, Figure 5.2C) in each of the four paired combinations, were tested on adult male and female *B. jarvisi* from two *B. jarvisi* lines and successfully amplified appropriately-sized products in males, with no amplification in females. These primers were subsequently used in the qRT-PCR experiments to sex individual embryos.

*Quantitative RT-PCR on sexed Bactrocera embryos*

Two fly lines were used for qRT-PCR. One was a laboratory stock of *B. jarvisi* (BJ). The second was a line carrying the complete *B. tryoni* autosome set in the presence of a *B. jarvisi* Y-chromosome (BTJ[Y]), obtained by hybridising virgin female *B. tryoni* and male *B. jarvisi*; followed by six generations of backcrossing hybrid males with virgin *B. tryoni* females.

Primers designed from the sequences described in the previous section were used for qRT-PCR. Sex-specific amplicons were targeted for *tra* and *doublesex*. At least one of each primer pair was designed across two adjacent exons to differentially target sex-specific transcripts and to minimise amplification of any carry-over DNA. The primers were tested for amplification on cDNA template from adult male and female RNA preparations from both experimental lines: *B. jarvisi* (BJ) and the hybrid line (BTJ[Y]), and on no-RT controls. Results demonstrated the sex-specificity of the amplified fragments. Amplification of a bulk preparation of unfertilised eggs from the BJ and the BT parental stock recovered some transcripts derived from maternal genes, deposited into the egg during oogenesis. These included *Sxl*, *tra-2* and female-specific *tra* and *dsx,* but it was not determined if the corresponding proteins were also present in the egg. Primers did not amplify male-specific transcripts of *tra* and *dsx*, nor the transcript of *slam* (non-sex-specific), from unfertilised eggs.

*Relative quantification of BJ and BTJ[Y] sex-determination genes*

*Sex-lethal* (S*xl*) is highly conserved between *B. tryoni* and *B. jarvisi*, and the same primer pair was used to amplify a non-sex-specific transcript from embryos of 1-8h after egg laying (AEL). No difference between male and female BJ embryos in the expression level of *Sxl* was noted, but a steady increase in transcript abundance was observed from 3-4h AEL (Figure 5.4A). Comparative transcript abundance in male and female BTJ[Y] embryos was likewise similar; however the increase in abundance of *Sxl* mRNA in both male and female occurred after 5h AEL (Figure 5.5A).

Expression levels of $tra^{F}$ followed the expected pattern based on our prior result that the female-specific form is present in unfertilised eggs and adult females but not in adult males. In BJ female embryos, $tra^{F}$ transcript abundance was fairly constant, and was similar to that found in male embryos until 5-6h, after which the quantity decreased until it was undetectable in 9h male embryos (Figure 5.4B). In contrast, the parity of male and female expression in BTJ[Y] embryos only began to differ at 8-9h (Figure 5.5B). The male form of *tra* ($tra^{M}$) was difficult to analyse because its amplification in single embryos was low and inconsistent. The male-specific form was occasionally detected in female embryos, consistent with findings in *C. capitata*, which showed the processing of pre-mRNA transcripts to the female form going through every possible combination of splicing order, including the male-specific form (Gabrieli *et al.* 2010). For this reason, $tra^{M}$ data were not included in the analysis.

The transcripts of *tra-2* did not differ between sexes, and did not appear to change in expression level over this time course (Figure 5.4C, 5.5C). It is possible that a concomitant breakdown of maternal transcripts with zygotic transcription of *tra*-2 would escape detection without a net difference in abundance. However, should *tra-2* transcripts be directly or indirectly targeted by *M*, some depression in transcript abundance would be expected.

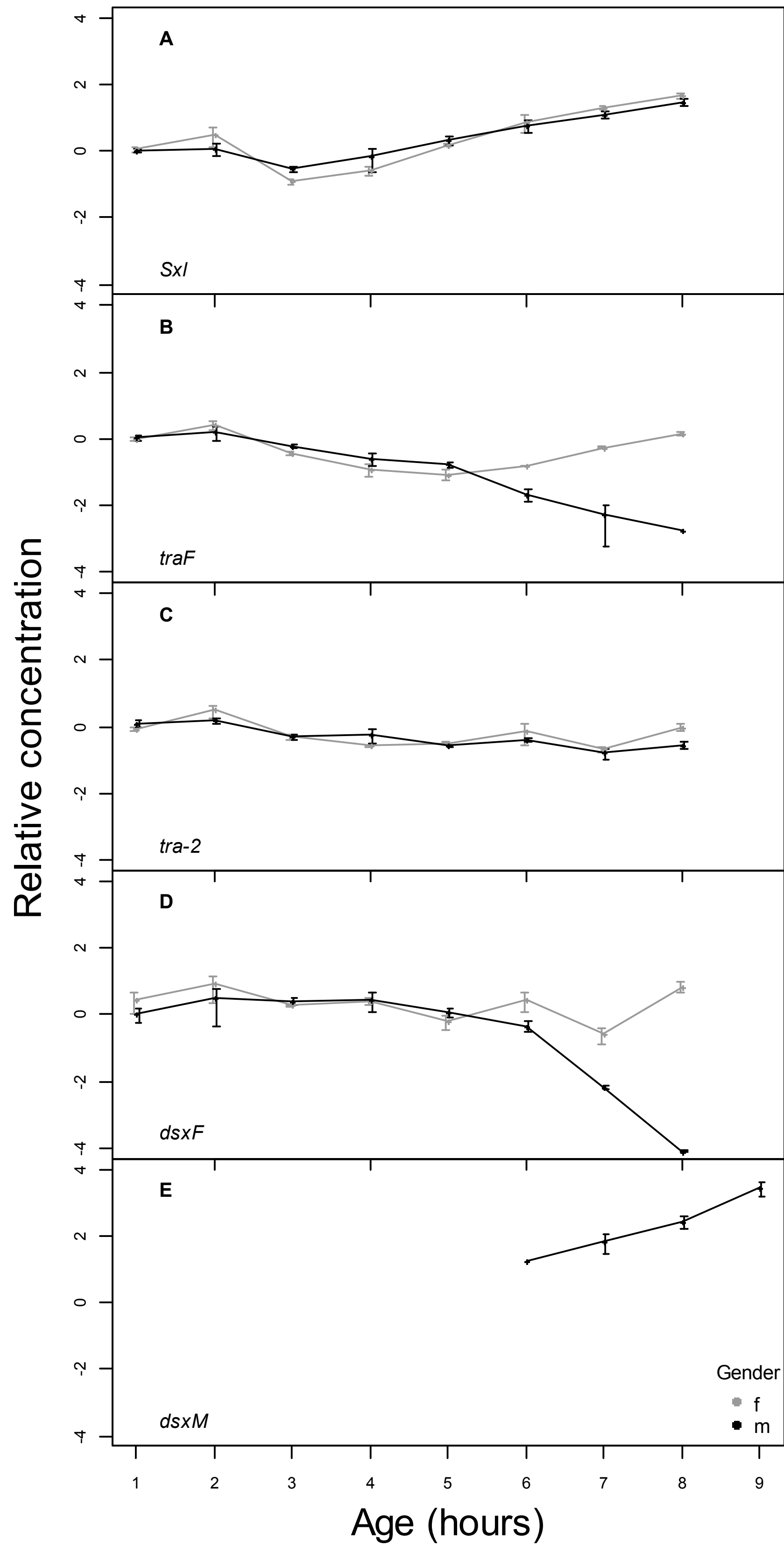



**Figure 5.4** qRT-PCR relative transcript abundance for BJ male (dark grey) and female (light grey) embryos aged from 1-9h. **(A)** common *Sxl* transcript, **(B)** female-specific spliced *tra* mRNA, **(C)** common *tra-2* transcript, **(D)** female-specific spliced *dsx* mRNA, and **(E)** male-specific spliced *dsx* mRNA.

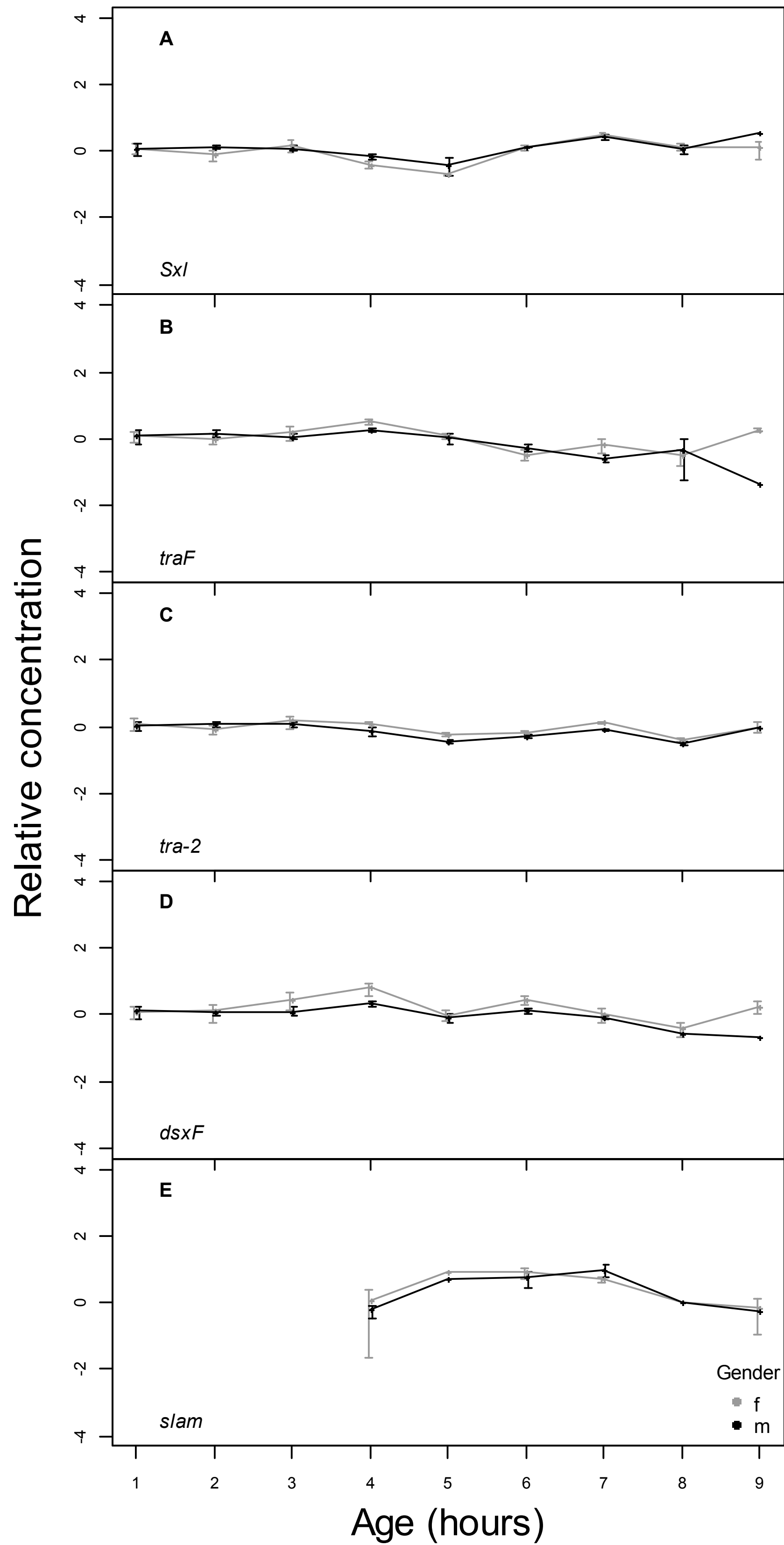



**Figure 5.5** qRT-PCR relative transcript abundance for BTJ[Y] male (dark grey) and female (light grey) embryos aged from 1-9h. **(A)** common *Sxl* transcript, **(B)** female-specific spliced *tra* mRNA, **(C)** common *tra-2* transcript, **(D)** female-specific spliced *dsx* mRNA, and **(E)** common *slam* transcript.

Female-specific *dsx* expression appeared relatively stable over the nine hours in female embryos, but exhibited a sharp drop in abundance between 6-7h in male BJ embryos (Figure 5.4D). The same pattern was found in BTJ[Y] female embryos (Figure 5.5D), but no sharp decline was apparent in males. Male-specific *dsx* was not detected in any embryo up to 6h AEL, but was detected in male BJ embryos between 6 and 7h AEL and increased in expression each hour to 9h (Figure 5.4E). Male BTJ[Y] embryos did not start expressing $dsx^{M}$ in the experimental time series (up to 9h), but it was detectable in 12h embryos by standard RT-PCR (Appendix D: Figure D.1).

*slam* expression was quantified only in BTJ[Y] embryos, and showed an increase in transcript levels between 3 and 4h AEL (Figure 5.5E). No *slam* was detected in any embryos of 1-3h, and the 4h embryos exhibited a wide variation in *slam* expression. This was likely due to natural variation in development, the 10min window of sampling and the rapid increase in transcript abundance during the fourth hour AEL. Standard RT-PCR on BJ embryos aged from 1-7h also failed to detect *slam* in 1-3h embryos, but was expressed from embryos 4h and older.

The difference in expression pattern between BJ and BTJ[Y] seemed to be only in the timing of early developmental changes in sex-determination gene expression, suggesting that development might be slower in BTJ[Y] than in BJ. The timing of developmental stages of BJ, the hybrid line BTJ[Y] and the parental BT line were compared with the reference data from the work of Anderson (1962), who comprehensively examined the developmental stages of *B. tryoni*. Phase contrast and fluorescence microscopic analyses were performed on embryonic stages visualised through propidium iodide staining, to compare the three lines and to link gene expression with developmental stages. BJ embryos were examined at 1h, 2h, 5h and 9h AEL; BTJ[Y] and the *B. tryoni* parental line (BT) were examined at 5h and 9h only (Figure 5.6). In both experimental lines, timing of development to syncytial blastoderm stage (at 5h) appeared to be very similar or identical – to each other, to the *B. tryoni* parent line, and to the published *B. tryoni* developmental time course (Anderson 1962). This equates to stage 4 (nuclear cycle 12) in *Drosophila* embryogenesis (Bownes 1975). However, in BJ 9h embryos, the pole cells were

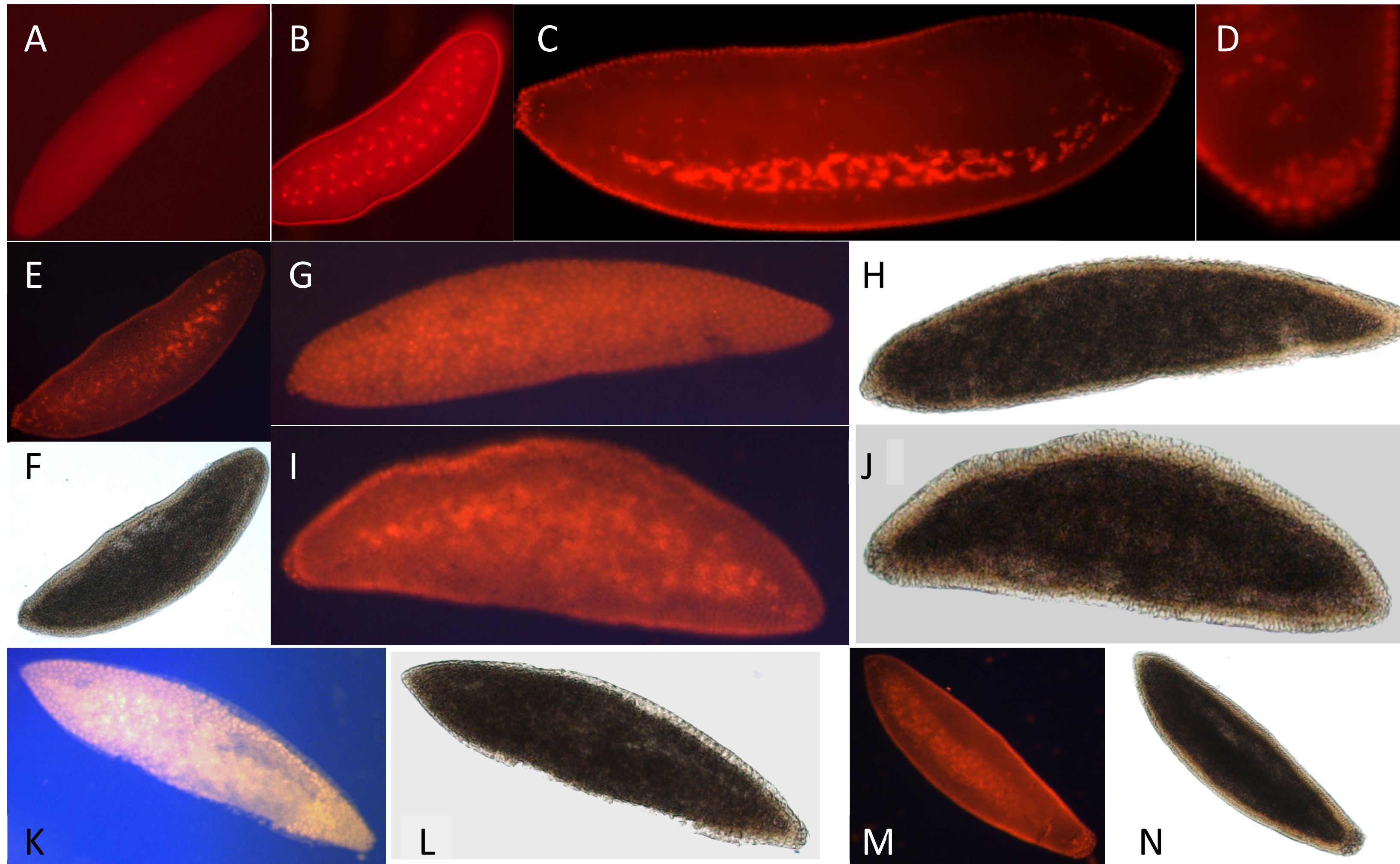



**Figure 5.6** *B. jarvisi* (A-F), *B. try*oni (G-J) and BTJ[Y] (K-N) embryos at different stages of early development (see next page for full legend)

**Figure 5.6** *B. jarvisi* **(A-F)**, *B. tryoni* **(G-J)** and BTJ[Y] **(K-N)** embryos at different stages of early development, stained with propidium iodide and viewed under fluorescence (A-E, G, I, K, M) or phase contrast (F, H, J, L, N). **(A)** *B. jarvisi* 1h AEL (stage 1), with four nuclei, **(B)** 2h AEL (nuclear cycle 7, stage 1) with 64 nuclei; **(C)** 5h AEL (nuclear cycle 12, stage 4) with surface nuclei, vitellophages along the midline and pole cells at posterior; **(D)** close view of pole cells separate from blastoderm nuclei in 5h embryo; **(E, F**) 9h embryo beginning to enter gastrulation with pole cells moving (stage 6). **(G, H)** *B. tryoni* 5h AEL embryo (cell cycle 12, stage 4); **(I, J)** 9h AEL embryo (cell cycle 14, stage 5). **(K, L)** BTJ[Y] 5h AEL embryo (cell cycle 12, stage 4); **(M, N)** 9h AEL (cell cycle 14, stage 5).

beginning to migrate dorsally, suggesting that the embryos were entering gastrulation, but no BTJ[Y] or *B. tryoni* embryos showed signs of gastrulation, although the full complement of 32 pole cells was visible and cellularisation was nearing completion (stage 5; Bownes 1975). The BT parent line cannot be sexed, but six, seven, six and four unsexed embryos of 7h, 8h, 9h and 12h AEL, respectively, were screened for male-specific *dsx* expression by RT-PCR, with only one out of four 12h embryos amplifying $dsx^{M}$. It is unlikely that all six or seven randomly sampled embryos are all female (Chi-squared, $p < 0.05$), so in BT, like in BTJ[Y], $dsx^{M}$ first appears between 9h and 12h AEL.

## 5.4 Discussion

Quantitative RT-PCR was used to compare the expression levels of a number of genes including those involved in sex determination and embryogenesis over a time course of early development in single embryos of two species of *Bactrocera* fruit flies. The Y-chromosome of *B. jarvisi* containing a genetic marker was introgressed into a *B. tryoni* line; examination of the cross-species effect of the *B. jarvisi M* on autosomal sex determination in *B. tryoni* demonstrated comparable mechanisms at work across these two species. However, the timing *of dsx*M expression in BTJ[Y] and the parental line BT, delayed as compared to BJ, indicates that the timing of the developmental regulation of sex-determination genes is species-specific, even though *M* is conserved. The critical period of sex-specific splicing of sex-determination genes was investigated independently in males and females of both species. Identifying this critical period is an important step towards identifying the likely time of action of the *M* through transcriptome sequencing of embryos of the relevant developmental stage.

*Conserved gene and transcript structure in tephritids*

The gene structure, features and transcript regulation of *Sxl*, *tra*, *tra*-2 and *dsx* in *B. tryoni* and *B. jarvisi* parallels other tephritids (*B. oleae*, *C. capitata* and *A. obliqua*) studied previously (see Table 1.1). Although *Sxl* is the key primary switch in *Drosophila*, exhibiting female-specific auto-regulation of splicing and translational repression (reviewed in Penalva and Sanchez 2003), *B. tryoni*, *B. jarvisi* and other tephritids do not generate sex-specific *Sxl* transcripts (Lagos *et al.* 2005, Saccone *et al.* 1998). The common *Sxl* transcripts are, however, highly conserved, present in the unfertilised egg and zygotically transcribed prior to the cellular blastoderm stage, suggesting an important role for *Sxl* in early development.

The *tra* auto-regulatory splicing mechanism proposed for tephritids is also supported by the presence of putative *cis*-acting TRA/TRA-2 binding sites and PRE sequences in the *B. tryoni* and *B. jarvisi tra* genes. These canonical sites are also located in the *Btdsx* mRNA (Shearman and Frommer 1998) and suggest that TRA and TRA-2 functional proteins are necessary at two major points in this pathway to ensure female development.

*Comparison of gene expression of tephritids*

Gene homologues of *Sxl*, *tra*, *tra-2* and *dsx*, displayed similar expression patterns in BJ and BTJ[Y]. The onset of zygotic transcription in BJ was shown to occur 3-4h AEL with the appearance of *slam* mRNA and the increase in mature mRNA of *Sxl* at this time. In BTJ[Y], despite no apparent change in *Sxl* transcripts, zygotic transcription was initiated at the same time, as evidenced by the detection of *slam* mRNA. In *D. melanogaster,* the maternal to zygotic transition (MZT), which involves two waves of zygotic transcription (Liang *et al.* 2008), heralds the increase in abundance of the *tra* transcript in females, and the subsequent translation of these into functional TRA to sufficient threshold levels, effecting the female-specific splicing of the *dsx* mRNA. Similarly, in BJ, the mRNA expression of zygotic sex-determination genes closely followed the onset of zygotic transcription, but is delayed by several hours in BTJ[Y]. The degradation of maternal *tra*$^{F}$ in males occurred at 5-6h, 8-9h and 6-7h in BJ, BTJ[Y] and *C. capitata* respectively.

For *tra-2*, transcript abundance did not vary over the 9 hour time course of the experiment, in either BJ or BTJ[Y]. Similarly, the expression of *Cctra-2* appeared to be stable over 10h of development in *C. capitata* (Gabrieli *et al.* 2010). An equal rate of transcription and degradation of *Bjtra*-2 and *Bttra*-2 might reflect the same appearance of stable expression. However, the *C. capitata* study found no splicing intermediates to suggest there was zygotic transcription (Gabrieli *et al.* 2010), and the *Drosophil*a model shows that maternal *tra-2* transcripts are stable (De Renzis *et al.* 2007). The absence of any substantial decrease in spliced *Bttra*-2, *Bjtra*-2 or *Cctra-2* transcript in males diminishes the possibility that degradation of *tra-2* transcripts, shown by RNAi to promote the male-specific splicing of *tra* and *dsx* (Salvemini *et al.* 2009), is the natural mechanism for interrupting the auto-regulation of *tra*.

Male-specific *Bjdsx* appeared at 7h in BJ males and marks the time at which sex is fully determined. This milestone was delayed in BTJ[Y] and BT, as no $dsx^{M}$ transcript was detected prior to 9h AEL, and in *C. capitata* none was detected until 10h AEL (Gabrieli *et al.* 2010). The continued detection of $Btdsx^{F}$ and $Bjdsx^{F}$ throughout early embryo development in female embryos differs from the complete degradation of *Ccdsx* until sex-specific forms are produced at 10h. This difference may reflect a difference in detection sensitivity, because qRT-PCR showed *dsx* transcripts were much less abundant than other maternally-expressed genes.

*Similarity of development times in BTJ[Y] and BT*

Anderson (1962) established that syncytial blastoderm in *B. tryoni* was completed at 5-6h, cellular blastoderm by 7.5h and gastrulation from 7.5h to 9h. We observed that under the same conditions, BJ appeared to progress at a similar rate through to the syncytial blastoderm stage, but gastrulation did not begin until approximately 9h AEL. Embryos of the hybrid line, BTJ[Y] and the BT parental line, also reached syncytial blastoderm at 5h, however no sign of gastrulation was seen in 9h embryos. It is not surprising that different species or different laboratory lines of a single species develop at different rates (Miyatake 1998, Souza *et al.* 1988), but it is notable that the early developmental stages progressed at the same rate, followed by a slowing of the rate of cellular development in BTJ[Y] and BT, compared to BJ. This

apparent delay in development occurred in embryos exposed to the same environment as the BJ parental stock line. Interestingly, the differentiation was also seen in the delayed expression of sex-determination genes in the BTJ[Y] line compared to BJ. It would appear that the regulation of genes expressed during early development differs between the two *Bactrocera* species, even though the *B. jarvisi M* is entirely compatible with *B. tryoni* autosomal sex-determination genes over a phylogenetic distance of approximately 30 million years (Krosch *et al.* 2012).

*Dominant Male Determiner*

In *B. jarvisi*, the time period between 3 and 6 hours AEL was identified as the time at which the *M* is likely to be active. Zygotic transcription was seen to have begun with the first detection of *slam* transcripts and increased abundance of *Sxl* transcripts. The direct effect of *M* was demonstrated by the coincident breakdown of the female-specific *tra* and *dsx* at 5h AEL in males, while abundance of both transcripts increased in females. This was followed at 6-8h AEL by the setting of *dsx* splicing in males to the male-specific form. Therefore, establishment of the sex-determined state preceded complete formation of the cellular blastoderm, as first noted in *M. domestica* (Hilfiker-Kleiner *et al.* 1993). However, in *D. melanogaster*, *dsx*$^{M}$ was not detected until gastrulation (Hempel and Oliver 2007). The time course of establishment of the sex-determined state in *B. tryoni* was later in the developmental program. It is not straightforward to relate this to the window of expression of *M*, because of the delayed breakdown of *tra*$^{F}$ in this species. There is a wider window of development time in which *M* might be active in *B. tryoni*, between 3-8h AEL.

The enduring effect of *M* activity is to ensure the absence of the functional TRA$^{F}$ isoform which in turn disrupts the establishment of the auto-regulatory *tra*$^{F}$ splicing. This could be achieved through transcriptional or translational repression, or inhibition of protein activity for either *tra* or *tra*-2. Saccone et al. (2002) proposed that *M* acts at the transcriptional or post-transcriptional level to inhibit either or both genes. In the present study, no reduction in *tra*-2 mRNA was observed in the key 3-6h AEL period. Abundance of *Bjtra* transcript in *B. jarvisi* does not differentiate the sexes until approximately 2h after zygotic transcription has been established and in *C. capitata*, *Cctra* transcription and splicing progress equally in male and female

embryos until 6h (Gabrieli *et al.* 2010), signifying that transcriptional repression is unlikely. In male *B. jarvisi* embryos, the slow reduction of *Bjtra* transcript to a non-detectable level over a three hour period indicates that the targeting of *tra* for degradation is not the likely role of *M*. Post-transcriptional splicing of *Cctra* is haphazard and inefficient, and it appears a similar mechanism may exist in *Bactrocera*, whereby the final female-specific transcript is produced after passing via the male-specific form.

Gabrieli et al. (2010) suggested that *M* resets *Cctra* splicing through suppression of $TRA^F$ protein activity. As splicing regulators of the SR type operate under a threshold dependency, a reduction in the amount of available TRA protein through direct inhibition, combined with inefficient splicing, would result in the determined male state. This mechanism is supported in *B. jarvisi* and *B. tryoni* because the effect of TRA or TRA-2 inhibition would influence the splicing of *tra* and *dsx* simultaneously. During the transition from 5-6h AEL in BJ and 8-9h in BTJ[Y], the female-specific mRNA expression of both *tra* and *dsx* is reduced. This occurs concurrently with the production of the male-specific form of *dsx*, and presumably *tra*.

## 5.5 Conclusions

The likely effect of *M* in male embryos is to block TRA or TRA-2 protein activity, so that female-specific splicing of zygotically transcribed *tra* mRNA is no longer facilitated. The ability of the *B. jarvisi M* to operate on autosomal sex-determination genes of both *B. jarvisi* and *B. tryoni* indicates functional conservation across species. The difference in the timing of developmental milestones, that we found in laboratory lines, but has also been noted in laboratory and field samples of other species, suggests that regulation of *M* must be co-ordinated with control of the autosomal sex-determination genes. If *M* is conserved across two *Bactrocera* species, it can also be examined across other *Bactrocera* species that can hybridise in the laboratory, and perhaps provide a generic application to the production of male-only genetic sexing strains. Characterising *M* may rest with transcriptomics focused on differentially screening male and female embryos sampled at the important developmental time. It would appear that the genetic expression of sex-determination

genes does correlate to cytological milestones, which may enable researchers to target appropriate time-periods for *M* expression in other tephritid species through linked analyses of cytology and gene expression.

## 5.6 Experimental procedures

*Fly Rearing*

Laboratory stocks of *B. jarvisi* (BJ) and *B. tryoni* (BT) were originally sourced from the Queensland Department of Agriculture, Fisheries and Forestry research facilities in Cairns, the NSW Department of Primary Industries in Gosford, respectively. These stocks were maintained at the University of NSW and University of Western Sydney, Australia, at constant 25°C, natural light and on artificial larval diet (Meats *et al.* 2004). Adult flies were fed protein one week after emergence and, between two and four weeks later, offspring were collected from mated females by inducing egg laying into a container of larval diet (Meats *et al.* 2004) covered with perforated parafilm.

*Bactrocera jarvisi* males are known to carry a mitochondrial *cytochrome B* DNA marker on the Y-chromosome (Shearman *et al.* 2010), which allows for discernment of the sex of embryos. In addition, by hybridisation with *B. tryoni* females, the same marker can be used to distinguish male from female *B. tryoni*. Hence, *B. tryoni* virgin females from the BT line were collected and caged with male BJ, their male offspring collected and backcrossed with virgin female BT. This backcross was repeated for six generations to make a line, BTJ[Y], with a predominant genetic background of *B. tryoni*, while carrying the introgressed Y-chromosome of *B. jarvisi*. The line was then maintained by full sib mating.

*Sex-determination genes*

Homologues of genes from the sex-determination pathway as described in *D. melanogaster* and *C. capitata* were targeted for this study. DNA sequence for *tra* and *tra-2* genes in *B. tryoni* and *B. jarvisi* was determined by using degenerate primers designed within conserved regions of amino acid alignments of

*D. melanogaster* and *C. capitata tra* (GenBank Accession No. NM079390 and AF434936 respectively) and *tra-2* (GenBank Accession No. NM057416 and EU999754 respectively). Primers for *Sxl* were designed from the *B. oleae Sxl* sequence (GenBank Accession No. AJ715415). Whole genome sequence for *B. tryoni* became available during the course of this research (Gilchrist *et al*. unpublished); *slam* was matched to the *C. capitata slam* sequence (GenBanAccession no. FG068639) and primers designed to exonic sequence. Primer sequences are detailed in Table D.1 (Appendix D).

DNA extraction of whole male and female adults of *B. tryoni* (BT) and *B. jarvisi* (BJ) followed the method described in Bennett and Frommer (1997). PCR using degenerate primers for *tra* and *tra-2* in various combinations (Appendix D: Table D.1, D.2) amplified products that were electrophoresed and excised from 1% agarose gels and purified with Wizard Gel Extraction Kit (Promega). Cloning, clone PCR and sequencing was performed as described in Chapter 2.

Total RNA was extracted from adult male and female *B. tryoni* and *B. jarvisi* separately, using Trizol (Invitrogen) reagent following the manufacturer's protocol. Reverse transcription was performed with SuperScript II reverse transcriptase (Invitrogen) using Oligo dT primer. Reverse transcription-PCR (RT-PCR) was employed to determine the *tra*, *tra*-2 and *Sxl* transcript exon sequence and structure in male and female flies, using specific primers designed from the genomic sequence of both *Bactrocera* species (primer sequences and PCR conditions are detailed in Appendix D: Table D1, D2). 3' and 5' RACE was performed principally as described in Frohman *et al*. (1988) using *tr*a and *tra-2* specific primers (Appendix D: Table D.1, D.2).

DNA and mRNA sequences were trimmed and aligned in Sequencher 4.0 (GeneCodes Corporation). A Y-chromosome pseudogene of the *tra-2* transcript, found only in male *B. jarvisi* genomic sequences, was aligned with other male and female genomic and mRNA sequences; primers were designed to amplify this sequence (Appendix D: Table D1) and were tested on DNA extracts from four males and four females from the same *B. jarvisi* line, and four individual males and females from a different collection of *B. jarvisi* (PCR protocol Appendix D: Table D.2).

*Embryo development times*

Significant milestones in early embryonic development were identified through both phase contrast and fluorescence microscopy. Embryos from all three lines were collected concurrently from three to four week old flies. Flies were supplied with an oviposition chamber containing apple juice and covered by perforated parafilm for a minimum of 1h; first-collected embryos were discarded and fresh chambers were supplied for 10min intervals. Embryos were carefully washed onto moist fabric and incubated at 25°C and 70% humidity to allow development to proceed to the desired time. At each appointed time, 15-20 embryos were taken, dechorionated by submersion in 1% sodium hypochlorite (Sigma) for 45-50s, and washed immediately in 0.3% Triton-X solution. Subsequent methanol fixation and propidium iodide (ICN) staining of embryos followed the procedure described for *Drosophila* embryos (Rothwell and Sullivan 2000). Embryos representing developmental times AEL of 1h, 2h, 5h, 7h and 9h were examined with the Zeiss Fluorescent microscope (Axio Scope.A1) and compared with images of *B. tryoni* (Anderson 1962) and *D. melanogaster* (Bownes 1975).

*Collection of material for qRT-PCR*

In separate experiments, BJ and BTJ[Y] females were induced to lay eggs onto parafilm moistened with apple juice over 10 minute periods. Minimal disruption to the embryos upon laying was critical for survival and development of the embryos to the appropriate time, which was an hourly time series from 1h to 9h. *Bactrocera* eggs are susceptible to desiccation and embryonic and larval mortality can be as high as 30% (J.L. Morrow, personal observation). Therefore, embryos were carefully transferred to moist cloth and incubated in a 70% relative humidity chamber at 25°C to continue development. To ensure optimal RNA quality, embryos were crushed with a microtube pestle in the presence of 100µL Trizol (Invitrogen) reagent before storage in the -80°C freezer.

*RNA and cDNA preparation*

Each embryo was individually processed for RNA extraction in preparation for quantitative RT-PCR, and DNA was isolated from these embryos for designation of sex. Prior to RNA extraction, 0.1ng of *Kanamycin* RNA (Promega) was added to each embryo sample. Extraction of RNA followed the method outlined in the Trizol protocol, scaled down to 100µL Trizol per embryo, unless Phase lock gel tubes (PLG; 5PRIME, Germany) were used, when an additional 40µL DEPC-$H_2O$ and 70µL chloroform (Sigma) were added to the tissue in Trizol before transferring to the PLG tubes. Centrifuge times and speeds were according to the PLG tube manufacturers recommendations, which was 5min at 12,000g at room temperature. The aqueous phase was pipetted into fresh tubes containing 5µg glycogen (Roche) and 70µL isopropanol (Sigma) and mixed. The samples were incubated at room temperature for 10min, sometimes with additional incubation at -20°C for several hours, before centrifugation at 12,000g at 4-8°C for 10min. The supernatant was decanted and the RNA was washed twice in 75% ethanol, air dried and resuspended in 15µL nuclease-free water. RNA quality was tested by examination of 260/280nm and 260/230nm ratios by Nanodrop spectrophotometry, and RNA integrity by non-denaturing agarose gel electrophoresis. RNA was stored at -80°C, or immediately DNase treated with Turbo DNA-free Kit (Ambion) following the manufacturer's protocol. Reverse transcription (RevertAid First Strand cDNA Synthesis Kit, Fermentas) was performed in 20µL reactions, using 9µL of template RNA and Oligo dT primer. Alien inhibitor RNA was added to each reaction at the recommended quantity to control for inhibition in the RT and subsequent qRT-PCR.

RNA from single adult male and female BJ and BTJ[Y] flies was isolated using the Trizol method as described in the manufacturer's protocol. RNA was then reverse transcribed with RevertAid. Controls lacking RevertAid enzyme were produced in parallel to allow determination of the specificity of qPCR primers to spliced RNA transcripts.

*DNA extraction and genetic sexing PCR*

In parallel to the RNA extraction, single embryos were sexed based on DNA. DNA is found in the interphase of the Trizol/Chloroform embryo mix after centrifugation. With the aqueous layer removed for RNA extraction, 100µL 1M Tris at pH 7.6-8.0 was added to the organic phase, mixed and incubated at room temperature for more than 15min. Following centrifugation at 12,000g for 10min, the aqueous layer containing DNA was transferred to a new tube, ethanol precipitated and resuspended in 10µL nuclease-free water. The sex of the embryos was determined by single or multiplex PCR targeting a common gene in males and females (*Sxl*) and the Y-specific pseudogene (*tra-2*, Appendix D: Table D.1, D.2). Products were gel electrophoresed to visualise presence or absence of the Y-specific amplicon. DNA extraction and subsequent sex assignment using this PCR-based method was optimised for the small quantities of DNA isolated from single embryos, but amplification sometimes failed due to the low number of nuclei, as few as 4 nuclei, in the youngest embryos (Figure 5.6). However, amplification of embryos over 3h of age was more reliable, and absence of amplification in these older embryos was interpreted as an indication that the embryo had not developed normally, and the samples were discarded. BTJ[Y] samples were also tested with *slam* to confirm the initiation of zygotic transcription in these embryos; all embryos at 4h or more of age amplified *slam* or were discarded.

*Target genes*

Sex-determination gene transcripts for *Sxl*, *tra*, *tra-2* and *dsx* were targeted for this study. To provide a marker for early zygotic transcription and to confirm that development had proceeded, primers designed from the *B. tryoni* genome (Gilchrist *et al*. unpublished) were used to amplify a small section of *slam* in both *B. tryoni* and *B. jarvisi* genomic DNA and cDNA.

Primers for qRT-PCR of target transcripts were designed using PrimerQuest, constrained to straddle introns to minimise DNA contamination across all loci, and also to be male or female specific for *tra* and *dsx* cDNAs. One endogenous reference gene, *leonardo* (*leo*) was chosen as primers had been designed and used in a previous

study in *B. tryoni* (Cook 2005), however it was not known if this gene would be expressed stably over the early developmental period being examined here. For this reason exogenous RNAs were also used: *Kanamycin* (*Kan*) RNA (Promega) was added at the beginning of the RNA extraction to normalise the RNA extraction of individual embryos and Alien qRT-PCR Inhibitor Alert RNA (Integrated Sciences) was added to the reverse transcription reaction to detect and account for inhibition in the reverse transcription and the subsequent PCR. Primers for *Kan* (An *et al.* 2002) produced a large amplicon of 793bp, Alien Inhibitor primers were proprietary and produced an amplicon of 239bp. All qPCR primer sequences, melting temperatures and amplicon sizes are listed in Table D.3 (Appendix D). PCR conditions for end-point PCR using qPCR primers were as decribed for RT-PCR (Appendix D: Table D.2).

*qPCR validation*

Prior to qPCR, validation of primer efficiency and specificity was carried out. The efficiency of primers was determined by qPCR on a 1/10 dilution series of the applicable PCR amplicons, generated from amplification of cDNA from male and female adult *Btj*. All efficiency (E) values except one fell between 0.95 and 1.07 (Appendix D: Table D.4). *Kanamycin* efficiency was low at 0.87, which is probably a consequence of the large size of the amplicon (763bp). All other primer pairs designed in this study amplified cDNA between 120bp and 189bp. As relative quantification was desired across the time course of samples and not across target genes, no efficiency correction was applied. Primer sets were designed across introns to eliminate annealing to genomic DNA and to be specific to male or female transcripts when appropriate. Primers were tested on no-RT control reactions and resulted in no product formation for any target gene. Specificity of primer sets designed to amplify male or female specific transcripts was tested on adult male and female cDNA. Male-specific primers did not amplify a *dsx* or *tra* product in adult females and female-specific primers did not amplify *dsx* or *tra* from adult males. All amplifications producing an amplicon were then tested by qPCR, and the melt curves were found to generate a single peak; subsequent gel electrophoresis confirmed a single product of expected size.

*qRT-PCR*

Quantitative RT-PCR of samples and standards were assembled in triplicate by a pipetting robot (CAS-1200, Corbett Research) and run in 100 well rotor-discs on the Rotorgene 6000 (Qiagen). Standards were made to test the efficiency of the primers. Each 20µL RT reaction was diluted 1 in 8.4 prior to PCR, to standardise the volume added to each reaction and minimise error. Reactions were 10µL, comprising 5µL SensiMix SYBR (Bioline, Sydney, Australia), 0.4µM each primer and 4.2µL of cDNA template (equivalent to 0.5µL of undiluted template). Cycling conditions were 95°C for 10min, then 45cycles of 95°C for 20s, 60°C for 30s and 72°C for 45s, followed by a disassociation cycle with incremental increase of 1°C from 60°C to 98°C every 5s.

Mean quantification cycle (Cq) values were calculated from the triplicate reactions, normalised against the geometric mean of three reference genes, *kanamycin*, Alien Inhibitor and *leo* (Vandesompele *et al.* 2002), and transformed to provide fold differences according to the formula $2^{-\Delta\Delta Cq}$ (Livak and Schmittgen 2001). Values were plotted in the R statistical environment (R Core Team 2012).

# Chapter 6

# Transcriptome sequencing of male and female *Bactrocera jarvisi* embryos

Morrow, J. L., Riegler, M., Gilchrist, A. S., Shearman, D. C. A. and Frommer, M. "Transcriptome sequencing of male and female *Bactrocera jarvisi* embryos" BMC Genomics (in preparation)

## 6.1 Abstract

Developing embryos are provided with maternal RNA transcripts and proteins, but transcription from the zygotic nuclei must be activated to control continuing development. Transcripts are generated at different stages of early development, and those involved in sex determination and cellularisation are some of the earliest to be activated. In *Bactrocera jarvisi* (Diptera: Tephritidae), transcription from the Y-chromosome is surmised to set in motion a cascade that determines male development, as part of the greater maternal to zygotic transition (MTZ). *Bactrocera jarvisi* embryos were collected over two pre-blastoderm time periods, 2-3h and 3-5h after egg laying and individually sexed using a Y-chromosome genetic marker. Sex-specific poly(A)$^+$ transcriptome was sequenced and the transcriptomes assembled *de novo.* Fifteen sex-determination gene homologues and two cellularisation gene homologues of *Drosophila melanogaster* (Diptera: Drosophilidae) were newly identified in *B. jarvisi*: *extra-macrochaetae* (*emc*) displayed a zygotic transcription profile contrary to the maternal expression in *D. melanogaster*; *sisterless A* (*sisA*) expression occurred very early; *slow as molasses* (*slam*) and *nullo* transcripts increased 80- and 17-fold respectively over time. No strong candidates for Y-chromosome transcripts were recovered from the poly(A)$^+$ fraction, perhaps indicating that such transcripts are not protein coding. These data contribute fundamental information to sex-determination research, and provide candidates for the sourcing of gene promoters for transgenic pest-management strategies.

## 6.2 Background

Early stages of embryonic development involve large changes to the RNA transcript profile, as maternal transcripts, deposited during oogenesis, are targeted for degradation, and activation of the zygotic genome takes place. In *Drosophila melanogaster,* egg activation is triggered by osmotic and physical stimulation, occurs independently of fertilisation and requires proteins such as SMAUG (SMG) and microRNAs to regulate degradation of maternal mRNAs (Chapter 5; Bushati *et al.* 2008, Tadros *et al.* 2007, Tadros and Lipshitz 2009). In *Drosophila*, at least 30% of transcripts have the distinctive expression profile of maternal transcripts and about two thirds of these decrease markedly over the first 6.5h of development (Arbeitman *et al.* 2002).

During these early stages, transcription from the zygotic genome must be initiated. Activation of zygotic transcription is controlled, in part, by the *Zelda* (*Zld*) protein interacting with TAGteam sites located upstream of genes targeted for early (pre-blastoderm) transcription (Liang *et al.* 2008, ten Bosch *et al.* 2006). Some of the earliest genes to be transcribed are involved in sex determination, such as *sisterless A* (*sisA*) during nuclear cycle 8 (Erickson and Cline 1993); cellularisation including *serendipity α* (*sryα*), *nullo* (*nullo*), *bottleneck* (*bnk*) and *slow as molasses* (*slam*) from cycle 11 (Lecuit *et al.* 2002, Rose and Wieschaus 1992, Schejter and Wieschaus 1993, Stein *et al.* 2002); and transcriptional activation, including transcripts and proteins that promote ongoing transcription and further enhance maternal mRNA degradation (Tadros and Lipshitz 2009).

As discussed in Chapters 1 and 5, the sex-determination pathway is highly conserved in Diptera, most notably at the terminal gene, *doublesex (dsx)*, and its upstream regulatory genes *transformer (tra)* and *transformer-2* (*tra-2;* see Table 1.1). In *D. melanogaster*, sex-specific splicing of *tra,* which generates the active TRA protein in females and a non-functional protein in males, is regulated by the *Sex-lethal* protein (SXL). Ongoing production of functional SXL occurs in females because an early, transiently-generated SXL is only produced in females in response to the primary signal, and mRNA transcribed from the late promoter is female-specifically spliced only when early SXL is present. The primary signal that leads to female-specific *Sxl* mRNA is transmitted by a combination of X-chromosome-linked

signal elements (XSE), namely *sisA*, *scute* (*sc*), *outstretched* (*os*) and *runt* (*run*), whose protein products are more concentrated in XX females than XY males (Salz and Erickson 2010). These gene products interact with maternal products *daughterless* (*da*), *hermaphrodite* (*her*), *extra-macrochaetae* (*emc*) and *groucho* (*gro*) and the zygotically-expressed autosomal gene *deadpan* (*dpn*) (reviewed in Penalva and Sanchez 2003; Chapter 1). Many of these genes have other molecular functions in development, and have been co-opted into a sex-determination regulatory role in *Drosophila* as the apex of the pathway diverged from other Diptera. *Sxl* is not sex-specifically spliced in other dipteran families such as the tephritids (Table 1.1).

Homologues of the principal genes *tra*, *tra2* and *dsx* have been sequenced in *Bactrocera* fruit flies (Chapter 5; Shearman and Frommer 1998), as has *Sxl*, which differs again from the *Drosophila* model by its maternal deposition in the egg. *Sxl* is also zygotically transcribed in the pre-blastoderm embryo in tephritid fruit flies (Chapter 5; Gabrieli *et al.* 2010), thus increasing its expression levels. EST libraries of the tephritid fruit fly *Ceratitis capitata* (Gomulski *et al.* 2008) have identified homologues for some of the genes involved in *Sxl* regulation, and their transcript expression analysed in unfertilised eggs and early developing embryos to provide a picture of the expression of these gene products in the critical stages of sex determination (Gabrieli *et al.* 2010).

In contrast to the well-studied model *Drosophila* species, the early stages when the sex-determination pathway is activated shows different modes of regulation in many non-drosophilid insects. In many dipterans, it appears that zygotic transcription from the Y-chromosome of male embryos enables the female-specific sex-determination gene transcripts, which are part of the maternal mRNA complement, to be replaced by male-specific transcripts, thus resetting cell memory. This putative Y-chromosome transcript is the *Dominant Male Determiner* (*M*), which has yet to be characterised, but the most likely target appears to be the TRA/TRA-2-based spliceosome, thereby effectively prohibiting the female-specific splicing of both *tra* and *dsx* pre-mRNA (Chapter 5). The expression of zygotic genes important in early developmental processes such as sex determination is expected to increase, perhaps rapidly, during embryonic stages prior to blastoderm formation.

*Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae) is the major fruit fly pest species in Australia, exhibiting a wide host fruit range and geographic distribution; and is the primary target of pest management strategies applied in Australia such as sterile insect technique (SIT; Andrewartha *et al.* 1967) and potentially the incompatible insect technique (IIT; Boller *et al.* 1976). *Bactrocera jarvisi* (Tryon) is a native Australian fruit fly found on the north and east coast of Australia, but due to its smaller host range is not as great a pest as *B. tryoni*. However, *B. jarvisi* has found great usefulness in research because of its tractable laboratory maintenance, ability to form fertile hybrids with the major pest species *B. tryoni*, and its possession of two Y-chromosome genetic markers, fundamental for discerning male and female embryos (Shearman *et al.* 2010; Chapter 5). These features were exploited to perform qRT-PCR expression analysis of genes involved in sex determination and cellularisation in single, sexed embryos from 1h to 9h after egg laying (AEL) in *B. jarvisi* and a *B. tryoni* line carrying the introgressed *B. jarvisi* Y-chromosome (Chapter 5). Moreover, the compatibility of the male determiner in *B. jarvisi* and *B. tryoni* anticipates strategies for optimising SIT and IIT in *B. tryoni*, such as the development of genetic-sexing (Schetelig and Handler 2012a), and embryonic lethality constructs (Schetelig and Handler 2012b). With germ-line transformation protocols established (Raphael *et al.* 2011), these strategies may be transferable between species, supported by access to whole genome assemblies for both *B. tryoni* and *B. jarvisi* (Gilchrist et al., unpublished).

The embryonic progression of *Bactrocera* species from fertilisation to cellularisation takes 7h. Expression analysis of *Sxl* and *slow as molasses* (*slam*) demonstrated that zygotic transcription has begun by 4h post fertilisation, and affects sex-specific transcripts from 5h (Chapter 5). Therefore, we chose to analyse poly(A)$^{+}$ transcriptome data from sexed *B. jarvisi* embryos of 2-3h and 3-5h AEL to investigate three aspects of early development. One focused on the expression of sex-determination genes important to the *Drosophila* sex-determination pathway. The aim was to get a broader insight into the expression profiles of genes that have been co-opted to regulate the sex-determination pathway upstream of the primary genes conserved across tephritids and drosophilids. Secondly, we searched the assembly for genes showing early zygotic expression in *D. melanogaster*, focusing on cellularisation genes, to identify homologues possessing early-acting promoters; and

thirdly we used differential expression analysis to search for male-specific transcripts, which are not maternally deposited. Such transcripts could include candidate Y-chromosome transcripts such as *M that* constitutes the beginning of the sex-determination cascade.

## 6.3 Methods

*Fly Rearing*

*Bactrocera jarvisi* (BJ) laboratory stock was originally sourced from the Queensland Department of Agriculture, Fisheries and Forestry, Australia, and maintained at University of New South Wales and then University of Western Sydney at constant 25°C, natural light and on artificial larval diet (Meats *et al.* 2004). Flies were fed protein in the form of yeast hydrolysate and sugar one week after emergence, and two to four weeks later gravid females were induced to lay eggs into larval diet covered by perforated parafilm.

*Embryo collection and RNA and DNA extraction*

Embryos of two different age ranges were selected for sequencing: 2-3h and 3-5h after egg laying. For each time period, male and female embryos needed to be separated based on a molecular marker on the Y-chromosome, as there are no morphological characters to distinguish early embryos. Each embryo was individually processed for RNA extraction. Further, DNA was isolated from these embryos for designation of sex. Embryos for RNA and DNA extraction were collected from *B. jarvisi* females by inducing egg laying through perforated parafilm coated with apple juice for a period of 1h, and then placed on moist fabric and maintained at 25°C for the appropriate time. Individual embryos were placed in microcentrifuge tubes, 100μL Trizol (Invitrogen) added and embryos crushed using microtube pestles (SSI). Samples were incubated at RT for several minutes and then stored at -80°C.

Extraction of RNA utilised TriSure (Bioline Australia), following the method described in Chapter 5 using PLG tubes (5PRIME, Germany). The quality of the

RNA was ascertained by non-denaturing gel electrophoresis of 1µL of each 15µL sample; samples with poor yield or degraded RNA were discarded, acceptable samples were pooled (see below) and stored at -80°C until required.

DNA is found in the interphase of the Trizol / Chloroform mix after centrifugation, this was extracted as described in Chapter 5. With the aqueous layer removed, 1M Tris at pH 7.6-8.0 was added to the organic phase, mixed and incubated for more than 15min at room temperature. Following centrifugation for 10min at 12,000g, the aqueous layer containing DNA was transferred to a new tube, ethanol precipitated and resuspended in 10µL nuclease-free water.

Multiplex PCR was performed on the DNA from each embryo with BjY2traB and BjY2traDrev primers to amplify the Y-chromosome fragment (227bp) and primers SxlRTFor1 and SxlRTRev1 to amplify from both sexes (280bp). Embryos were designated as male by the presence of both bands and female if the Y-chromosome band was absent and the 280bp *Sxl* band was produced. When multiplex PCR failed, single target PCR with primers BjY2traA and BjY2traDrev (311bp amplicon) and *Sxl* (as above) was carried out. PCR conditions are shown in Table E.1 (Appendix E). Many samples failed to amplify with any set of primers probably due to the low copy number; these were discarded.

RNA preparations from the two time points, following sex designation, were pooled to create two male samples at 3-5h AEL (26 and 23 embryos), two female samples at 3-5h AEL (27 and 17 embryos), two male samples at 2-3h AEL (28 and 24 embryos) and two female samples at 2-3h AEL (12 and 24 embryos). Pooled samples were ethanol precipitated and washed six times in 75% ethanol and resuspended in 30µL DEPC-$H_2O$. Quality of the RNA preparations was ascertained by Nanodrop spectrophotometry, Qubit DNA and gel electrophoresis (Appendix E: Table E.2). Poly(A)$^+$ selection, reverse transcription and library construction followed the Illumina RNA-Seq protocol and were performed at the Next Generation Sequencing Facility, Hawkesbury Institute for the Environment, UWS (Richmond, Australia). Eight paired-end libraries (300bp), comprising the two female samples and two male samples at both 2-3h AEL and 3-5h AEL time points, were sequenced on two lanes of the Illumina HiSeq 1000. The sequencing output was demultiplexed and provided as two paired fastq files for each sample.

*Transcriptome assembly and annotation*

These raw data were quality trimmed and filtered using CLC Genomics Workbench ver.6 (CLCbio) with parameters allowing 2 ambiguous nucleotides, minimum read length of 50 nucleotides and error probability limit of 0.05 applied to Phred quality scores according to the modified Mott trimming algorithm (see CLC Genomics Workbench v6 manual). *De novo* assembly of reads from all eight samples combined was performed with CLC Genomics using default parameters (herein referred to as "CLC assembly"). Similarly, Trinity (http://trinityrnaseq.sourceforge.net; Grabherr *et al.* 2011, Haas *et al.* 2013) was employed for *de novo* assembly from approximately half of the reads from all eight samples (every second pair from each sample was extracted and used for the assembly due to limits of RAM) using default parameters (herein referred to as "Trinity *de novo* assembly"). TransDecoder, a utility in the Trinity package that searches for long open reading frames and, from this group, selects transcripts that are probable coding regions, was applied to the Trinity *de novo* assembly. This subset of contigs ("Trinity CDS assembly") was used as the query in blastx searches of the NCBI non-redundant (nr) protein database followed by gene ontology assignment, implemented in Blast2GO (Conesa *et al.* 2005). Within the CLC workbench, the entire CLC assembly was queried against the NCBI nr protein database (downloaded October 2013) using blastx (cut-off E-value 1E-3).

The CLC assembly was interrogated for twenty-eight genes involved in sex-determination processes in *Drosophila* spp. (GO:0007530) and four genes involved in cellularisation. First, the genes were sought amongst the top BLAST hits and accepted if the E-value was <1E-3. For those genes that did not find a match to *B. jarvisi* sequences in the CLC assembly, *D. melanogaster* and *C. capitata* sequences from GenBank were aligned in Mega 5.05 (Tamura *et al.* 2011), and a highly conserved section of the *C. capitata* gene was used in a motif search of the CLC assembly contigs. The contigs with matches were then realigned with the *D. melanogaster* or *C. capitata* sequences in Mega 5.05 to confirm homology and to determine the length of the ORF. Patterns of expression of known sex-determination genes (*Sxl*, *tra*, *tra-2* and *dsx*; see Chapter 5) from qRT-PCR were compared to the RNA-Seq output. The cellularisation gene *slam* was shown to be transcribed between three to four hours (AEL) by qRT-PCR analysis (Chapter 5) and the *sisA* transcript was identified in this study (see Results). In *D. melanogaster*, *slam* is expressed in

nuclear cycle 11 and *sisA* is transcribed in nuclear cycle 8. The expression levels of these genes were assessed to determine if zygotic transcription was proceeding in the 2-3h embryos, and to determine an appropriate RPKM (Reads Per Kilobase of exon model per Million mapped reads) value threshold, over which expression values are likely to indicate maternally-derived transcripts.

*Differential expression*

RNA-Seq analysis was performed on the CLC assembly within CLC Genomics Workbench. The trimmed paired reads from each sample were mapped to the CLC assembly using default parameters which include minimum length (0.9) and similarity (0.8) fractions and expression values reported as RPKM. Differential expression experiments were run on fourteen combinations of the eight samples across time and gender and the RPKM values were quantile normalised and fold changes calculated. The proportions-based (Baggerley's) test was applied to normalised expression values, differences were selected when the false discovery rate (FDR)-corrected $p < 0.001$. To investigate zygotic expression, differential expression output was filtered to select transcripts up-regulated in 3-5h versus 2-3h embryos and male versus female embryos. Selections were filtered further by restricting the mean normalised RPKM values in the baseline sample to a variable number, based on assessment of *slam* and *sisA* transcript expression.

### 6.4 Results

A developmental time series of *B. jarvisi* transcript levels of important sex-determination genes, *Sxl*, *tra*, *tra-2* and *dsx*, enabled us to choose two time periods of development to target for transcriptome analysis (Chapter 5). The period 2-3h AEL was selected because it fell in the phase when zygotic transcription was just beginning, while still containing enough nuclei to confidently determine the sex of the embryo; and 3-5h AEL to cover the phase when zygotic transcription is proceeding, including the time when *M* is expected to be active.

*Transcriptome assembly*

The Illumina RNA-Seq output for the eight *B. jarvisi* samples is described in Table 6.1, with samples representing a range of 53 million to 112 million reads prior to quality trimming and filtering. The filtered sequencing reads from all eight samples were assembled *de novo* by two methods. CLC Genomics assembled 61,223 contigs of average size 699bp (Table 6.2). The majority of contigs were between 200 and 500bp (64.6%), with 34.2% over 500bp in length. Homology searches using blastx (NCBI nr protein database, October 2013) returned 23,518 sequences (E-value < 1E-3), including 260 (1.1%) sequences matching *Bactrocera* species; 16,246 (69.1%) matching sequences from *Ceratitis* spp., mostly from *C. capitata* (16,209); and 2,596 (11.0%) sequences matching *Drosophila* species. Of those sequences homologous to *Bactrocera*, the majority corresponded to *B. dorsalis* (156), *B. oleae* (64) and *B. tryoni* (29).

Trinity *de novo* assembly (http://trinityrnaseq.sourceforge.net) was applied to approximately half of the paired trimmed reads due to constraints on RAM. This produced 63,848 contigs of average length 1,203bp (Table 6.3). This dataset was too large to annotate using Blast2GO, therefore a subset of 22,085 transcripts was extracted, using TransDecoder, these were uploaded into Blast2GO (Conesa *et al.* 2005) for BLAST matching and gene ontology (GO). BLAST found matches for 19,687 sequences (89.1%), the top hits distributed primarily among *Drosophila* spp., most prominently *D. virilis*, *D. mojavensis* and *D. melanogaster* (Figure 6.1).

Gene ontology (GO) analysis procured 14,004 annotated transcripts (69.7%), categorised into 36 functional groups within the classes molecular function, cellular component and biological process (Figure 6.2). The largest representations were in “binding” and “catalytic activity” (molecular function), “cell” (cellular component) and “cellular process” and “single-organism process” (biological process). While 2,291 sequences mapped to reproduction (GO:0000003) and 5,446 sequences mapped to developmental processes (GO:0032502), no sequences mapped to sex determination (GO:0007530) or developmental processes involved in reproduction (GO:0003006).

**Table 6.1** Results of trimming and filtering poor quality reads in CLC Genomics Workbench.

| Sample ID | Paired reads (fastq files) | Number of reads | Average length | Number of reads after trim | Percentage trimmed | Average length after trim |
|---|---|---|---|---|---|---|
| BJ1 | JM1_CAGATC_L007_R1_001/JM1_CAGATC_L007_R2_001 | 112,478,286 | 101 | 110,748,851 | 98.46% | 99.9 |
| BJ2 | JM2_ATCACG_L005_R1_001/JM2_ATCACG_L005_R2_001 | 63,790,670 | 101 | 61,969,621 | 97.15% | 98.7 |
| BJ3 | JM3_CGATGT_L005_R1_001/JM3_CGATGT_L005_R2_001 | 55,946,780 | 101 | 54,337,396 | 97.12% | 98.7 |
| BJ4 | JM4_TTAGGC_L005_R1_001/JM4_TTAGGC_L005_R2_001 | 92,808,680 | 101 | 89,984,188 | 96.96% | 98.6 |
| BJ5 | JM5_ACTTGA_L007_R1_001/JM5_ACTTGA_L007_R2_001 | 98,766,172 | 101 | 97,235,566 | 98.45% | 99.9 |
| BJ6 | JM6_TGACCA_L005_R1_001/JM6_TGACCA_L005_R2_001 | 74,893,404 | 101 | 72,254,387 | 96.48% | 98.6 |
| BJ7 | JM8B_GCCAAT_L005_R1_001/JM8B_GCCAAT_L005_R2_001 | 82,963,712 | 101 | 80,449,239 | 96.97% | 98.7 |
| BJ8 | JM8_ACAGTG_L005_R1_001/JM8_ACAGTG_L005_R2_001 | 53,199,536 | 101 | 51,563,810 | 96.93% | 98.6 |

Settings: Removal of low quality sequence (limit=0.05); removal of ambiguous nucleotides (maximal 2 nt); removal of sequences of minimum length 50 nt.

**Table 6.2** CLC *de novo* assembly of transcripts from all eight libraries combined.

| Summary Statistics | Count | Average length | Total bases | Nucleotide distribution | Count | Frequency |
|---|---|---|---|---|---|---|
| Reads | 604,738,912 | 99.24 | 60,014,654,938 | Adenine (A) | 13,031,025 | 30.40% |
| Matched | 476,830,889 | 99.24 | 47,320,571,906 | Cytosine (C) | 7,923,401 | 18.50% |
| Not matched | 127,908,023 | 99.24 | 12,694,083,032 | Guanine (G) | 7,925,335 | 18.50% |
| Reads in pairs | 418,126,768 | 154.99 | | Thymine (T) | 13,018,385 | 30.40% |
| Broken paired reads | 58,704,121 | 98.54 | | Any nucleotide (N) | 899,616 | 2.10% |

| Contig measurements | Length |
|---|---|
| Count | 61,223 |
| N75 | 434 |
| N50 | 1,138 |
| N25 | 2,697 |
| Minimum contig (bp) | 114 |
| Maximum contig (bp) | 17,400 |
| Average (bp) | 699 |
| Total (bp) | 42,797,762 |

**Table 6.3** Trinity *de novo* assembly of sequences from all eight *B.jarvisi* libraries.

| Summary Statistics | Count |
|---|---|
| Number of contigs | 63 848 |
| Average length of contigs | 1 203bp |
| Number of reads | 309 000 000 |
| N80 | 848bp |
| N50 | 2 887bp |
| N20 | 5 693bp |
| minimum | 201bp |
| maximum | 18 709bp |
| sum | 7.68E+07 |
| n:200 | 63 848 |
| n:N50 | 7 769 |

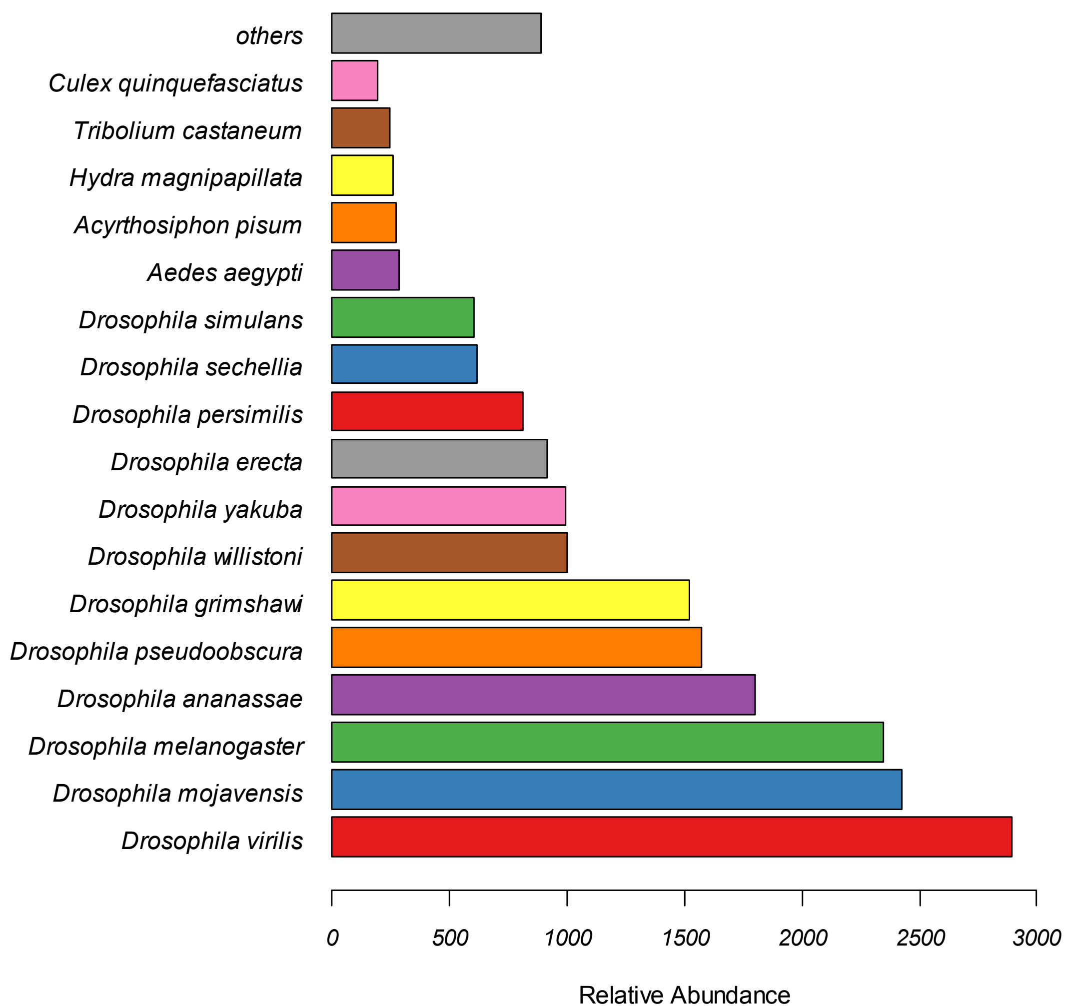



**Figure 6.1** Species distribution of BLAST2GO top matches against the Trinity CDS assembly (i.e. 22,085 of the total 63,848 contigs that are likely coding sequences, extracted by TransDecoder). Over 89% of the sequences matched *Drosophila* species.

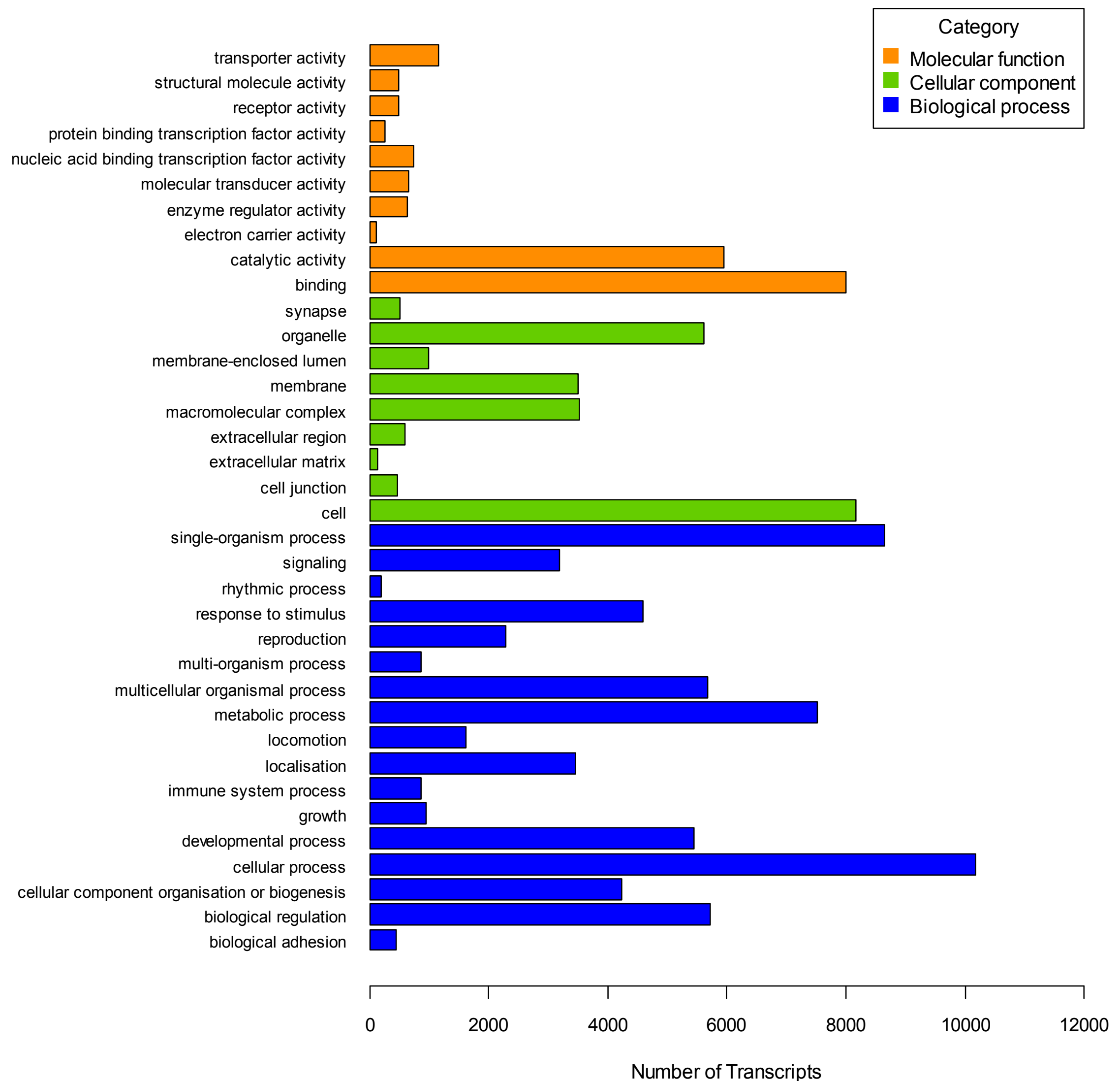



**Figure 6.2** Classification of the gene ontology (GO) terms for the *B. jarvisi* transcriptome (implemented in BLAST2GO on the Trinity CDS assembly). 14,004 (69.7%) contigs were classified into 36 functional groups.

*RNA expression analysis and validation*

Within the RNA-Seq module of CLC Genomics, each trimmed library was mapped to the CLC assembly and levels of expression were recorded as normalised RPKM values. Validation of the assembly and mapping was performed. Principal components analysis of the eight samples mapped to the CLC assembly showed samples clustering according to age, except that two samples, BJ1[male 3-5h] and BJ5[male 2-3h], clustered together and not with their replicates (Figure 6.3).

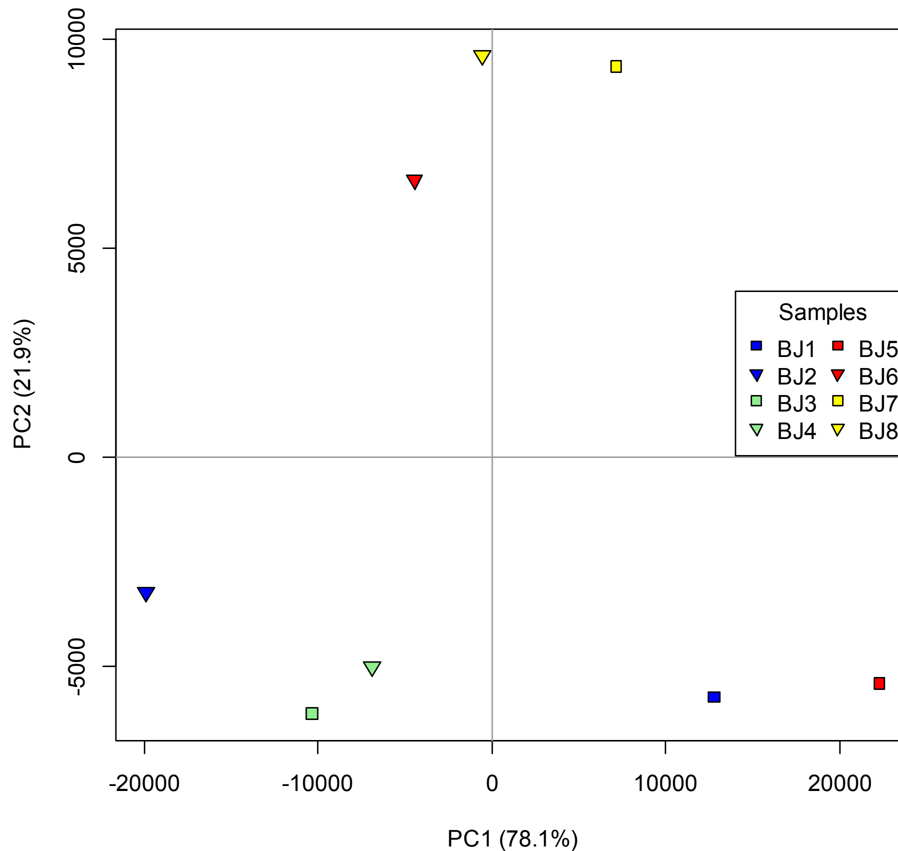



**Figure 6.3** Principal co-ordinates analysis of the eight transcriptome samples mapped to the CLC assembly. Samples are male3-5h (blue), female 3-5h (green), male 2-3h (red) and female 2-3h (yellow).

To investigate this discrepancy, the comparative expression levels over the two time points for the sex-determination genes *Sxl*, *tra*, *tra-2* and *dsx* and the cellularisation gene *slam* (Chapter 5) served as controls to validate the RPKM values (Tables 6.4 and 6.5). As there was no difference between male and female embryos in expression pattern for *Sxl*, *tra-2* and female-specific *tra* and *dsx* over the first 5h, differential expression across the two developmental periods, irrespective of gender, was examined. However, the results for *Sxl* and *slam* would be most useful because both exhibit an increase in transcript abundance over this time course. *Sxl* RPKM values indicated that BJ1[male 3-5h] showed the lower expression levels equivalent to the 2-3h samples, rather than the higher levels of the other 3-5h samples (Table 6.4, Appendix E: Table E.3).

**Table 6.4** Expression profiles of the 19 sex-determination genes identified in the *B. jarvisi* CLC assembly. Expression in each sample BJ1 – BJ8 is reported as the normalised RPKM value from the comparison of early samples (BJ5-8) to late samples (BJ1-4). Also refer to Appendix E: Table E.3.

| | | | | "late" samples | | | | "early" samples | | | | Significant up or down-regulation Comparison | | Validated expression profile# | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | male | | female | | male | | female | | | | | | |
| | contig no. | contig size | ORF size* | BJ1 | BJ2 | BJ3 | BJ4 | BJ5 | BJ6 | BJ7 | BJ8 | late vs early | late vs early (exclude BJ1 and BJ5) | Bj | Cc | Dm |
| *doublesex* | 22304 | 1064 | 235aa * | 0.5 | 0.04 | 0.05 | 0.06 | 0.52 | 0.22 | 0.14 | 0.19 | ns | ns | M | M | Z |
| | 40032 | 808 | 95aa * | 0.48 | 0.07 | 0.22 | 0.005 | 0.16 | 0.16 | 0.17 | 0.29 | ns | ns | | | |
| *daughterless* | 16154 | 4053 | 710aa | 1.55 | 5.1 | 4.61 | 4.96 | 3.19 | 3.64 | 10.69 | 10.77 | ns | ns | | | M |
| *deadpan* | 20953 | 3365 | 574aa | 0.99 | 12.12 | 12.18 | 27.25 | 0.29 | 0.24 | 0.29 | 1.25 | ns | 0.02 | | Z | Z |
| *extra-macrochaetae* | 10611 | 1230 | 265aa | 7.59 | 132.81 | 133.21 | 224.54 | 0.98 | 0.56 | 0.5 | 2.84 | ns | <0.001 | | | M |
| *female lethal d* | 2854 | 2694 | 466aa * | 18.57 | 41.45 | 46.02 | 38.89 | 20.08 | 14.5 | 12.73 | 11.42 | 0.04 | <0.001 | | M | M |
| *fruitless* | 4880 | 1862 | 524aa * | 20.26 | 31.36 | 28.2 | 36.67 | 30.27 | 24.33 | 26.21 | 25.78 | ns | | | | |
| | 29049 | 3175 | 675aa | 0.6 | 0.21 | 0.39 | 0.09 | 0.77 | 0.65 | 0.77 | 0.44 | ns | | | | |
| *groucho* | 3466 | 4604 | 726aa | 14.16 | 53.64 | 61.73 | 98.47 | 16.65 | 7.92 | 13.93 | 12.86 | ns | <0.001 | | M | M |
| *hopscotch* | 753 | 4443 | 1174aa | 42.66 | 15.81 | 20.55 | 24.5 | 56.97 | 21.64 | 24.55 | 19.65 | ns | ns | | | |
| *Mes-4* | 5091 | 5397 | 1557aa | 35.36 | 12.04 | 19.58 | 18.02 | 60.46 | 18.74 | 24.01 | 17.26 | ns | ns | | | |
| *ovarian tumor* | 3340 | 3198 | 653aa * | 28 | 24.46 | 33.02 | 20.67 | 39.37 | 27.68 | 75.67 | 55.11 | ns | ns | | | |
| | 5442 | 201 | 68aa * | 14.92 | 13.17 | 15.3 | 10.64 | 13.66 | 13.78 | 43.13 | 24.74 | ns | ns | | | |
| | 5443 | 201 | 68aa * | 2.66 | 0.1 | 0.83 | 1.01 | 5.38 | 0.57 | 0.65 | 0.39 | ns | ns | | | |
| | 8947 | 1757 | 459aa * | 14.59 | 15.73 | 17.52 | 10.75 | 21.49 | 15.6 | 46.3 | 33.38 | ns | ns | | | |
| | 10161 | 300 | 101aa * | 3.54 | 0.31 | 0 | 1.38 | 8.42 | 0 | 1.03 | 1.02 | ns | ns | | | |
| *ovo* | 21175 | 255 | 86aa * | 1.25 | 0.08 | 0.59 | 0.1 | 3.36 | 2.81 | 8.01 | 14.67 | ns | ns | | | |
| | 22105 | 317 | 34aa * | 4.06 | 0.91 | 0.17 | 0.35 | 6.46 | 8.76 | 39.76 | 44.79 | ns | ns | | | |
| | 23602 | 317 | 34aa * | 0.71 | 0.17 | 0 | 0.08 | 2.65 | 2.73 | 15.39 | 3.45 | ns | ns | | | |
| *runt* | 13344 | 1521 | 268aa * | 10 | 56.83 | 57.58 | 29.75 | 0.53 | 0.28 | 0 | 4.54 | ns | <0.001 | | | Z |
| | 1493 | 840 | 204aa * | 13.45 | 74.67 | 62.21 | 47.82 | 0.92 | 0.2 | 0.12 | 6.49 | 0.02 | <0.001 | | | |
| | 1494 | 229 | 61 aa * | 6.56 | 36.06 | 24.92 | 18.43 | 0.8 | 0.3 | 0 | 2.65 | 0.05 | <0.001 | | | |

| | | | | "late" samples | | | | "early" samples | | | | Significant up or down-regulation Comparison | | Validated expression profile^ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | male | | female | | male | | female | | | | | | |
| | contig no. | contig size | ORF size* | BJ1 | BJ2 | BJ3 | BJ4 | BJ5 | BJ6 | BJ7 | BJ8 | late vs early | late vs early (exclude BJ1 and BJ5) | Bj | Cc | Dm |
| *sans fille* | 7153 | 1421 | 150aa | 46.68 | 68.71 | 58.91 | 54.45 | 38.3 | 85.76 | 79.64 | 91.13 | ns | 0.006 | | M | M |
| *scute* | 57147 | 340 | 114aa * | 0.04 | 0.26 | 0.01 | 0.07 | 0.08 | 0.09 | 0.09 | 0.09 | ns | ns | | | Z |
| | 50842 | 514 | 172aa * | 0 | 0.04 | 0.08 | 0.15 | 0.05 | 0.06 | 0.05 | 0.19 | ns | ns | | | Z |
| *Sex-lethal* | 2265 | 1667 | 444aa * | 24.32 | 41.51 | 46.32 | 69.96 | 25.97 | 22.39 | 25.77 | 24.24 | ns | ns | M | M | Z |
| *sisterless A* | 2508 | 264 | 89aa * | 80.91 | 167.57 | 190.76 | 261.16 | 60.92 | 38.36 | 26.51 | 34.29 | 0.02 | <0.001 | | Z | Z |
| *transformer* | 4523 | 3261 | 327aa * | 30.6 | 18.43 | 18.49 | 21.87 | 30.32 | 25.07 | 20.58 | 20.46 | ns | ns | M | M | Z |
| | 13992 | 392 | 48aa * | 12.87 | 10.04 | 8.28 | 13.79 | 12.92 | 9.74 | 6.75 | 6.8 | ns | ns | | | |
| *transformer-2* | 138 | 3165 | 251aa | 51.33 | 39.42 | 39.5 | 42.65 | 51.08 | 46.66 | 44.53 | 45.35 | ns | ns | M | M | M |
| *virilizer* | 2897 | 994 | 308aa * | 35.7 | 32.32 | 37.51 | 34.89 | 26.7 | 38.12 | 41.35 | 43.84 | ns | ns | | | Z |
| | 4042 | 392 | 131aa * | 31.71 | 30.87 | 33.41 | 31.12 | 27.84 | 35.35 | 37.09 | 40.38 | ns | ns | | | |
| | 4266 | 3049 | 1002aa * | 21.87 | 21.43 | 22.68 | 18.79 | 21.82 | 22.82 | 28.76 | 24.18 | ns | ns | | | |
| | 4267 | 1131 | 352aa * | 19.84 | 18.28 | 21.63 | 17.2 | 16.01 | 19.61 | 21.41 | 19.91 | ns | ns | | | |
| | 5176 | 234 | 79aa * | 15.63 | 9.71 | 15.54 | 3 | 9.88 | 10.23 | 6.31 | 12.1 | ns | ns | | | |
| | 5177 | 234 | 79aa * | 16.33 | 25.95 | 15.03 | 35.46 | 13.29 | 28.43 | 35.73 | 28.82 | ns | ns | | | |
| | 7162 | 268 | 90aa * | 12.02 | 14.66 | 8.39 | 18.8 | 12.09 | 19.44 | 21.31 | 20.34 | ns | ns | | | |
| | 7163 | 268 | 90aa * | 12.64 | 6.76 | 11.11 | 3.29 | 9.59 | 6.68 | 5.09 | 8.67 | ns | ns | | | |
| | 12608 | 218 | 57aa * | 16.39 | 13.53 | 17.67 | 15.67 | 11.41 | 20.53 | 16.22 | 15.51 | ns | ns | | | |
| | 16540 | 309 | 104aa * | 3.38 | 6.83 | 5.95 | 1.83 | 2.82 | 5.94 | 3.44 | 4.06 | ns | ns | | | |
| | 16541 | 310 | 104aa * | 14.66 | 5.98 | 15.41 | 1.93 | 9.42 | 5.5 | 5.26 | 8.38 | ns | ns | | | |

*Partial ORF

^Experimentally validated in *B. jarvisi* (Bj) see Chapter 5; *C. capitata* (Cc) and *D. melanogaster* (Dm)

# maternal (M) or zygotic (Z) expression

**Table 6.5** *Bactrocera jarvisi* cellularisation gene homologues. Cellularisation gene homologues were found by performing a motif search of the CLC assembly using *D.melanogaster* (*bottleneck* and *nullo*) and *C. capitata* (*slow as molasses* and *serendipity α*) sequences. No match was found for *bottleneck*. RPKM values for *nullo* and *slow as molasses* increase in older embryos, displaying expression profiles consistent with *D.melanogaster* cellularisation genes.

| | | | | "late" samples | | | | "early" samples | | | | Significant up or down-regulation Comparison | | Validated expression profile | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | male | | female | | male | | female | | | | | | |
| | contig no. | contig size | ORF size* | BJ1 | BJ2 | BJ3 | BJ4 | BJ5 | BJ6 | BJ7 | BJ8 | late vs early | late vs early (exclude BJ1 and BJ5) | Bj | Cc | Dm |
| *bottleneck* | | | no match | | | | | | | | | | | | | Z |
| *nullo* | 9415 | 202 | 68aa * | 6.92 | 132.64 | 96.17 | 41.85 | 1.07 | 2.18 | 0.29 | 13.2 | ns | ns | | | Z |
| | 6453 | 764 | 112aa * | 11.33 | 207.53 | 160.09 | 61.73 | 1.49 | 3.19 | 0.13 | 21.93 | ns | 0.03 | | | Z |
| | 14101 | 1026 | 66aa * | 2.64 | 38.95 | 31.36 | 15.9 | 0.4 | 0.33 | 0.15 | 12.09 | ns | | | | |
| | 14102 | 593 | 107aa * | 1.27 | 31.92 | 25.75 | 3.92 | 0 | 1.06 | 0 | 0.76 | ns | | | | |
| | 14117 | 1443 | 87aa * | 2.08 | 48.19 | 30.99 | 8.38 | 0.28 | 0.45 | 0 | 0.38 | ns | | | | |
| | 15641 | 455 | 22aa * | 0.77 | 15.8 | 10.66 | 5.5 | 0.3 | 0.31 | 0 | 3.6 | ns | | | | |
| *serendipity α* | 9105 | 5007 | 615aa | 12.89 | 4.41 | 3.74 | 4.58 | 13.75 | 11.64 | 10.63 | 9.82 | ns | ns | | | Z |
| *slow as molasses* | 3824 | 5366 | 1568aa | 38 | 485.68 | 458.67 | 323.24 | 2.04 | 0.87 | 0.35 | 14 | ns | <0.001 | Z | Z | Z |

* Partial ORF

The RPKM values for *slam*, a zygotically transcribed gene validated by RT-PCR (Chapter 5) were low in BJ1[male 3-5h] (38 RPKM) compared to the other three 3-5h samples (323-485 RPKM), and low in three 2-3h samples (0.35-2.04 RPKM) and relatively high in the fourth 2-3h sample, BJ8[female 2-3h] (14 RPKM). The difference in RPKM values over time periods was not significant for *slam*, unless BJ1[male 3-5h] alone or BJ1[male 3-5h] and BJ5[male 2-3h] were excluded (FDR < 0.001; Table 6.5, Appendix E: Table E.3). This phenomenon was also observed for *sisA*, *run*, *gro*, *female lethal d* (*fl(2)d*) and *emc* (see sex-determination gene analysis, below), where each gene exhibited increased transcription in the other three 3-5h embryos (and significant up-regulation following exclusion of BJ1[male 3-5h] and BJ5[male 2-3h]; FDR<0.001). The *dpn* transcript, zygotically expressed in *C. capitata* embryos, also appeared to be expressed this way, that is, very low RPKM values in 2-3h samples (0.24-1.25) and significantly higher in 3-5h embryos, except for BJ1[male 3-5h] (0.99 RPKM). Likewise, *emc* appears to be transcribed early in the zygote, contrasting with its maternal deposition in *Drosophila*, but in BJ1[male 3-5h] *emc* is expressed 17 to 29-fold lower than the other three 3-5h samples. Furthermore, *sisA*, one of the earliest zygotic transcripts in *D. melanogaster* and *C. capitata*, (but not yet validated by RT-PCR in *B. jarvisi*) was already showing high expression in 2-3h samples.

Taken together, these results suggest two features of the samples that may cause discrepancies: firstly, the age range of one hour in the 2-3h samples and 2 hours in 3-5h samples introduces variability, especially in the expression of early or rapidly increasing transcripts, such as *slam*, but may also be seen in early zygotic activity of *emc*, *dpn* and *run* (for example, sample BJ8 appears to comprise more developed embryos). Secondly, the expression levels of BJ1[male 3-5h] transcripts appear to result either from a preponderance of embryos at the earlier end of the range, or from a sample in which many embryos did not survive to develop for the full 3-5h, thus producing a weaker increase in expression levels for those genes undergoing zygotic transcription, or a higher apparent expression value for maternal transcripts undergoing degradation in older embryos. Therefore, differential analysis was performed both with and without the BJ1[male 3-5h] and BJ5[male 2-3h] samples, employing 14 combinations to identify changes in transcript expression. The number of contigs differentially expressed (FDR<0.001), displayed in Table 6.6, shows

clearly that when time comparisons are undertaken, greater differences are recovered in the absence of samples BJ1[male 3-5h] and BJ5[male 2-3h], with BJ1 responsible for the greater impact. As a result, and to maintain consistency, the following analyses were performed without BJ1 and BJ5.

*Sex-determination gene expression*

No sex-determination genes (GO:0007530) were detected in the Trinity CDS assembly. The CLC assembly was examined for the 28 *Drosophila* genes categorised as sex-determination. Using the blastx annotation of the CLC assembly and additional motif searches, the contigs were screened and 19 genes secured a contig match (Table 6.4, Appendix E: Table E.3). The matching contigs were verified by alignment with the database nucleotide sequence and the ORF determined. This process also helped to assess the quality of assembly through identification of complete or partial open reading frames (ORFs). Only nine transcripts contained complete ORFs in a single contig, two more contigs covered over 90% of the ORF (for *Sxl* and *fl(2)d*); the remaining eight genes were traversed by two to 11 contigs which covered more than 90% of the ORF (Table 6.4). The limitations of assembly appeared to be the result of different variants or alleles, but it also may be due to insufficiently stringent quality control.

Four of these genes: *Sxl*, *tra-2*, female-specific *tra* and *dsx*; are recognised maternal transcripts in *B. jarvisi* (Table 6.4) that exhibited either no substantial change over the 2-5h period or, for *Sxl*, increased expression between 3 and 4h (Chapter 5). RNA-seq of 2-3h and 3-5h embryos detected high levels of these transcripts, except for *dsx* where low levels (normalised RPKM<1) were detected in all samples. No significant change in expression from one time sample to the next was recorded. Other sex-determination genes detected by blast homology and expressed at high levels in 2-3h embryos are also possibly maternally derived. These include *da*, *fl(2)d*, *fruitless*F (*fru*), *gro*, *hopscotch* (*hop*), *Mes-4* (*Mes-4*), *ovarian tumour* (*otu*), *sans fille* (*snf*) and *virilizer* (*vir*) (Table 6.4). Other transcripts that are first transcribed in the zygote include *dpn* and *emc*, which have low RPKM in 2-3h samples (0.24-1.25 and 0.5 – 2.84 respectively).

**Table 6.6** Summary of differential expression experiments performed in CLC Genomics Workbench. Expression values were normalised and significant difference ascertained by Baggerley's test, with FDR-corrected p value cutoff at 0.1%. Contigs were selected for up-regulation in males over females and 3-5h embryos over 2-3h embryos using normalised fold-change.

| | Experiment | Samples[1] | No. of differentially expressed contigs (FDR p<0.001) | No. up-regulated in males and/or older embryos | Additional filters | No. of transcripts of interest |
|---|---|---|---|---|---|---|
| | **Male/Female comparisons** | | | | | |
| A | Late only (3-5h) | 1,2 vs 3,4 | 2 | 0 | | 0 |
| B | Late only (3-5h), excl. BJ1 | 2 vs 3,4 | 21 | 9 | †mean normalised RPKM of female samples <5 | 5 |
| C | Early only (2-3h) | 5,6 vs 7,8 | 73 | 4 | †mean normalised RPKM of female samples <5 | 0 |
| D | Early only (2-3h), excl. BJ5 | 6 vs 7,8 | 142 | 46 | †mean normalised RPKM of female samples <5 | 4 |
| E | Both (3-5h and 2-3h) | 1,2,5,6 vs 3,4,7,8 | 0 | 0 | | |
| F | Both (3-5h and 2-3h) excl. BJ1 and BJ5 | 2,6 vs 3,4,7,8 | 0 | 0 | | |
| G | Both (3-5h and 2-3h) excl. BJ1 | 2,5,6 vs 3,4,7,8 | 0 | 0 | | |
| | **Time comparisons** | | | | | |
| H | Both (Male and female) | 1,2,3,4 vs 5,6,7,8 | 283 | 86 | | |
| I | Both (Male and female) excl. BJ1 and BJ5 | 2,3,4 vs 6,7,8 | 2313 | 1213 | #mean normalised RPKM of 2-3h samples <50 | 1051* |
| J | Both (Male and female) excl. BJ1 | 2,3,4 vs 5,6,7,8 | 1846 | 1128 | #mean normalised RPKM of 2-3h samples <50 | 1019 |
| K | Male only | 1,2 vs 5,6 | 14 | 2 | | |
| L | Male only excl. BJ1 and BJ5 | 2 vs 6 | 1321 | 827 | #mean normalised RPKM of 2-3h samples <50 | 673* |
| M | Male only excl. BJ2 and BJ6 | 1 vs 5 | 183 | 66 | | |
| N | Female only | 3,4 vs 7,8 | 1980 | 862* | | |
| | **Intersections** | | | | | |
| | I ∩ L | | | | | 341 |
| | (I ∩ L) - (I ∩ L ∩ N) | | | | | 105 |

[1] Samples BJ1- BJ8 are abbreviated to 1-8

†Allowed extra expression (RPKM<5) in female samples to cover any accidental inclusion of male embryos

# Determined cutoff to be 20% higher than the mean expression, in 2-3h embryos, of *sisA*, a transcript expressed early in *D.melanogaster*

*Subsets of I, L and N used in intersections

*Cellularisation genes*

Some genes involved in cellularisation in *Drosophila* species are activated early in embryonic development. The sequence of *slam* was found in the CLC assembly, with a mean RPKM of 4.3 in 2-3h embryos and 326 in 3-5h embryos (Table 6.5). No *slam* was detected in young embryos by qRT-PCR (Chapter 5), but transcript levels rapidly increased over the next few hours of development. Although *slam* was detected here in one sample at a level higher than expected (RPKM=14), the other three 2-3h samples had low levels (RPKM<1), indicating that *slam* is expressed at low levels earlier than 3h and then increases in expression substantially (80-fold) and rapidly. Three additional cellularisation genes *sryα*, *nullo* and *bnk* were sought through blast annotation and motif searches of both CLC and Trinity assemblies. A potential match was found for *nullo* with 29.6% identity across 102 amino acids with *Drosophila simulans nullo* (Acc. No. DSY44733) and 57% identity over a conserved stretch of 28 amino acids. Other than *nullo* homologues from ten *Drosophila* species, there are no records of *nullo* in the NCBI database. This sequence displayed a similar expression profile as *slam*, with a 17-fold increase over time (Table 6.5). Neither *sryα* nor *bnk* were detected through the blast annotation, but a motif search of a 50bp fragment of the *C. capitata* sryα gene (Acc. No. FJ460703) identified a single contig with a partial ORF of 615 amino acids, exhibiting 61.6% and 33.5% identity with *C. capitata* and *D. melanogaster sryα* respectively. The expression profile was inconclusive, but transcript levels diminished as the embryos aged (Table 6.5).

*Data mining for differentially expressed transcripts*

For male / female comparisons, no differentially expressed transcripts were detected that adhered to criteria of low or zero female expression with higher expression in males (Table 6.6, experiments A, C, E, F, G), except for B and D, identifying five and four transcripts (Table 6.7). Annotation revealed contaminant fungal, bacterial, and ribosomal RNA sequences in eight of them. One 220bp sequence fragment was expressed 23-fold higher in male embryos at 2-3h of age: this sequence matched the *B. dorsalis even-skipped* (*eve*) pair-rule gene. However, inclusion of male and female data from 3-5h, showing higher expression in females as time proceeds, but a reduction in the male samples over time, does not lend support for this transcript to

**Table 6.7** Differential expression results showing significantly up-regulated transcripts in males over females in 3-5h samples (Experiment B) and 2-3h samples (Experiment D).

<table>
<tr><th colspan="3">Experiment B</th><th colspan="3">Normalized expression values (RPKM)</th><th></th><th></th><th></th></tr>
<tr><th colspan="3">Late only (3-5h), excl. BJ1</th><th>male</th><th colspan="2">female</th><th colspan="3">Top BLAST hit</th></tr>
<tr><th>Feature ID</th><th>Gene length</th><th>Fold Change (normalized values)</th><th>BJ2</th><th>BJ3</th><th>BJ4</th><th>NCBI Accession No.</th><th>E-value</th><th>Gene / Source</th></tr>
<tr><td>contig 22537</td><td>1546</td><td>-109.63</td><td>185.79</td><td>0.1</td><td>3.29</td><td>EQB46020</td><td>7.85E-55</td><td>hypothetical protein CGLO_15009 [Colletotrichum gloeosporioides Cg-14] (fungal pathogen of plants)</td></tr>
<tr><td>contig 26828</td><td>407</td><td>-90.35</td><td>145.56</td><td>0</td><td>3.22</td><td>ELT94824</td><td>2.06E-40</td><td>hypothetical protein CAPTEDRAFT_122939, partial [Capitella teleta] (Annelid worm)</td></tr>
<tr><td>contig 30721</td><td>521</td><td>-44.23</td><td>68.22</td><td>0.39</td><td>2.7</td><td>EEH16716</td><td>2.00E-28</td><td>senescence-associated protein [Paracoccidioides brasiliensis Pb03] >gi|225678434|gb|EEH16718.1| (fungi)</td></tr>
<tr><td>contig 47470</td><td>424</td><td>-100.05</td><td>111.08</td><td>0.12</td><td>2.1</td><td>XP_662849</td><td>3.93E-26</td><td>hypothetical protein AN5245.2 [Aspergillus nidulans FGSC A4] >gi|40742997|gb|EAA62187.1| (fungi)</td></tr>
<tr><td>contig 49000</td><td>222</td><td>-39.38</td><td>131.19</td><td>0.96</td><td>5.71</td><td>EHK21457</td><td>3.99E-41</td><td>hypothetical protein TRIVIDRAFT_53758, partial [Trichoderma virens Gv29-8] (fungi)</td></tr>
<tr><td colspan="9">Experiment D<br>Early only (2-3h), excl. BJ5</td></tr>
<tr><td></td><td></td><td></td><td>BJ6</td><td>BJ7</td><td>BJ8</td><td></td><td></td><td></td></tr>
<tr><td>contig 31948</td><td>667</td><td>-387.12</td><td>64.8</td><td>0.1</td><td>0.23</td><td>CAJ30045</td><td>7.17E-28</td><td>conserved hypothetical protein [Magnetospirillum gryphiswaldense MSR-1] (bacteria)</td></tr>
<tr><td>contig 3465</td><td>234</td><td>-20.53</td><td>29.21</td><td>0.57</td><td>2.28</td><td>XP_004522331</td><td>3.36E-45</td><td>PREDICTED: 60S acidic ribosomal protein P0-like [Ceratitis capitata] >gi|20139848|sp|Q9U3U0.1|</td></tr>
<tr><td>contig 41160</td><td>210</td><td>-70.11</td><td>37.85</td><td>0.33</td><td>0.75</td><td>WP_006574328</td><td>1.06E-16</td><td>hypothetical protein [Pseudoflavonifractor capillosus] >gi|150270408|gb|EDM97731.1(bacteria)</td></tr>
<tr><td>contig 567</td><td>220</td><td>-23.73</td><td>33.44</td><td>2.82</td><td>0</td><td>ACN91520</td><td>2.50E-16</td><td>eve [Bactrocera dorsalis]</td></tr>
</table>

be considered a good Y-chromosome candidate. To broaden the search for transcripts up-regulated in males, further comparisons were made to distinguish these transcripts; the most useful experiments were I, L and N (Table 6.6). The differential expression between the 2-3h embryos and the 3-5h embryos of both genders (I) revealed significant change of expression in 2313 transcripts, of which 1213 were up-regulated in older embryos (FDR-corrected $p<0.001$) indicating zygotic transcription. The experiment of BJ2[male 3-5h] and BJ6[male 2-3h] (L) was not replicated, but was used to find the overlap with the sequences from I ($I \cap L$). A subset of sequences expressed at no more than 50 RPKM in the 2-3h samples was used to eliminate transcripts that are 20% more abundant than *sisA*, a conservative estimate to minimise the elimination of candidate transcripts that are highly expressed in the 2-3h embryos. As a result, 341 transcripts common to both sets (1051 and 673 sequences respectively) were extracted. To reduce the inclusion of false positives, experiment N, highlighting up-regulated sequences in older female embryos, was added to exclude those sequences also transcribed strongly in female embryos. The resulting 105 sequences had top blast hits to 53 *C. capitata* sequences, 9 *Drosophila*, four transposon-related sequences from other arthropods, and 39 undetermined sequences (E-value > 1E-3 or no match retrieved; Appendix E: Table E.4, E.5).

### 6.5 Discussion

By utilising next-generation sequencing technology, we sequenced the transcriptome of *B. jarvisi* embryos during two developmental periods prior to blastoderm formation. Notably, by exploiting a molecular marker on the Y-chromosome of *B. jarvisi* males, this dynamic developmental period was examined independently in male and female embryos to distinguish transcripts differentially regulated between the sexes and over time. The identification, regulation and sequence of sex-determination genes in *B. jarvisi* was the primary focus of this study and was approached in two ways: through homology-based identification of validated sex-determination genes from *Drosophila* and homologues from other tephritids including *C. capitata*; and expression-based studies highlighting up- or down-regulation over time and differences between the sexes.

Only 32 *B. jarvisi* records are currently stored in NCBI databases (November 2013): nine mitochondrial, 20 transposons and three ITS and ribosomal RNA sequences. Chapter 5 reports sequences for *Sxl*, *tra* and *tra-2*. The paucity of molecular information for this fruit fly has limited the avenues of investigation into molecular processes involved in early development. The collection of transcriptome data are a substantial addition to the knowledge base of this fruit fly species and other closely related *Bactrocera* species.

*Different roles for sex-determination genes*

Sex-specific differences in expression levels of the four XSEs found in *Drosophila* were not observed for these genes in *B. jarvisi*. The four genes, *sisA*, *sc*, *os* and *run* are transcribed from both X chromosomes in *Drosophila* females, to communicate the female signal to *Sxl*. Bearing in mind the non-sex-specificity of *Sxl* in tephritids, we identified *B. jarvisi* homologues of *sisA*, *sc* and *run*, but not *os*, expressed in the early developmental stages; *sisA* and *run* increased 6- and 26-fold respectively over the two periods but did not differ in male and female embryos, and *sc* did not change over time. The X and Y chromosomes of *B. tryoni* are highly heterochromatic and the genes located on the *Drosophila* X-chromosome are autosomal in tephritids (Zhao *et al.* 1998), with no evidence that these genes play a role in sex determination. Of the four *Drosophila* XSEs, *os* is expressed latest (cycle 13) which may account for its absence in *B. jarvisi* embryos, or more likely the *B. jarvisi* homologue was not sufficiently similar to the known *Drosophila os* sequences. However, while *sisA* is expressed during cycle 8 and *sc* during cycle 9 of *Drosophila* embryos (Erickson and Cline 1993), in *B. jarvisi*, *sisA* is highly expressed even in the younger embryos and may be displaying a rapid increase as development proceeds, and *sc* is expressed very low in both time periods. Transcripts of *sisA* are not detected in *C. capitata* unfertilised eggs, hence it is likely that *B. jarvisi* also conforms to this expression profile, and that zygotic transcription of *sisA* is well underway between 2 and 3h AEL, however this must be confirmed experimentally. An autosomal sex-determination gene in *Drosophila* expressed in the zygote is *dpn*, which complies with this pattern in *B. jarvisi*, and exhibits more than 28-fold increase in transcript levels from 2-3h to 3-5h AEL.

Maternal transcripts in *Drosophila* include *her*, *da*, *gro* and *emc*. Homologous sequence for *her* was not detected in *B. jarvisi* embryos, but *da* and *gro* were detected in 2-3h embryos and *gro* exhibited a significant 6-fold increase over time. Confirmation is needed for the maternal input of these two genes, but expression patterns indicate that they are maternal transcripts. This status was also demonstrated for *gro* in *C. capitata* (Gabrieli *et al.* 2010). The expression of *emc* differs markedly from that expected for a maternal transcript – withvery low expression detected in 2-3h embryos, and increasing by 125-fold in 3-5h embryos. Unconventional gene expression profiles compared to *Drosophila* may indicate the loss or gain of functional roles in the developing embryo. In *D. melanogaster*, EMC is a negative regulator of *Sxl* via the formation of heterodimers with DA or SC, inhibiting the binding of these proteins to the *Sxl* promoter, and it is possible that this gene represents a gain-of-function in this species.

*Early zygotic transcripts as a source of promoters and enhancers for transgenic pest control strategies*

A partial sequence for a putative homologue of the *D. melanogaster* gene *nullo* was identified by blast annotation, coupled with an expression profile emulating a known cellularisation gene, *slam*, which is also expressed from nuclear cycle 11 in *D. melanogaster* (Lecuit *et al.* 2002, Rose and Wieschaus 1992). The *sryα* transcript however, undetected in *C. capitata* prior to cellularisation (Schetelig *et al.* 2009) but reported as a maternal contribution by Gabrieli et al. (2010), was found in *B. jarvisi* 2-3h embryos but had diminished levels in 3-5h embryos. Both *slam* and *nullo*, in addition to *sisA,* which also appears to be expressed early in *B. jarvisi*, provide candidate promoters for use in transgenic constructs. The promoter / enhancer regions of the *sryα* and *slam* genes have been used to create embryonic lethality systems in *C. capitata* (Schetelig *et al.* 2009) and the *Anastrepha suspensa sryα* homologue has been successfully applied to the construction of embryonic lethal (Schetelig and Handler 2012b) and female-specific lethal transgenes (Schetelig and Handler 2012a).

*Up-regulated transcripts in male embryos*

We were unable to confidently discern poly(A)$^+$ RNA that was more abundant in males and very low in females, and therefore potentially transcribed from the Y-chromosome. The Y-chromosome is highly heterochomatic, therefore few genes may be transcribed, furthermore, those transcripts may not be collected in the poly(A)$^+$ fraction of the transcriptome. At the time this experiment was performed, no whole genome sequence was available for *B. jarvisi*, and the sequencing of the poly(A)$^+$ fraction of mRNA was an essential step for *de novo* assembly, providing numerous novel coding sequences with homology to known *C. capitata* and *Drosophila* spp. genes. These reads have added valuable sequence data to assist in the assembly of the *B. jarvisi* genome (Gilchrist et al. unpublished). Y-specific transcripts may yet be found by obtaining sufficient coverage to identify the Y-chromosome sequences by subtraction of the genome of female *B. jarvisi* from the male genome. Such an endeavour is underway, thus transcripts from the male samples could be mapped onto the Y-chromosome and expression analysis repeated to determine up-regulated Y-specific transcripts as the embryos age.

## 6.6 Conclusion

Several experiments are warranted to extend, improve and confirm the results presented here. Firstly, confirmation of the maternal or zygotic transcription profile of some genes by gene-specific RT-PCR, or more comprehensively, by RNA-seq on unfertilised eggs is necessary for two genes. For *emc*, the methods here indicate it is zygotically transcribed during the 3-5h period, differentiating it from the mode of expression in *D. melanogaster*. *sisA* appears to have high zygotic expression during the 2-3h period and provides potential targets, along with *slam* and *nullo*, for promoter / enhancer sequences that may be used in transgenes to drive genes of interest. This includes inducing female lethality for male-only release cohorts for optimising SIT or IIT (Schetelig and Handler 2012a); or transgenes that drive embryonic lethality in crosses with wild flies, eliminating the need for sterilising the parental flies and thus avoiding the negative fitness effects (Schetelig *et al.* 2009, Schetelig and Handler 2012b).

The next step is to utilise whole genome sequence for *B. jarvisi*. Genome-guided transcriptome assembly was not feasible in this time frame, however applying this method may uncover transcripts that are differentially expressed or alternatively spliced or located on the Y-chromosome. In addition, identification of all transcripts will require sequencing of the non-coding fraction of the transcriptome, as many regulatory RNAs belong to that category (Mattick and Makunin 2006). The related pest fruit fly, *B. tryoni*, will soon have an annotated genome available, and so homology to the genes discovered here will allow valuable comparisons between the two species, which are reproductively compatible. This will be immediately useful for the identification of cellularisation gene promoters, such as *slam* and *nullo*, for transgenic-based improvements to pest management strategies such as sterile insect technique (SIT).

# Chapter 7

# General Discussion

## 7.1 Introduction

Australia is home to over 80 species of *Bactrocera* fruit flies (Drew 1989), some of which are highly invasive pest species that must be monitored and controlled in cost-effective and environmentally sustainable ways. *Bactrocera tryoni* is the most significant pest because of its broad endemic distribution and large host fruit range. *Bactrocera neohumeralis*, a sibling species , is as damaging to fruit over a narrower geographic range, and *B. jarvisi* is causing damage to cultivated fruit in its native range in northern Australia (Hancock *et al.* 2000). Invasive pest species of south-east Asia, such as the *B. dorsalis* group with previous incursions in Australia, and *C. capitata*, a worldwide pest species also established in Western Australia (Dominiak and Daniels 2012), prove very costly due to fruit damage and disinfestation measures. They require ongoing quarantine controls to either prevent them from spreading or to keep them out of Australia entirely.

Australian fruit fly research has focused on its most significant and endemic pest species, *B. tryoni,* and, because of its invasive potential, research on *B. tryoni* has mainly been restricted to Australia. A large body of work has described the biology (reviewed in Fletcher 1987) and ecology (reviewed in Clarke *et al.* 2011), with particular attention to various pest control methods (reviewed in Dominiak and Ekman 2013), but limited focus on biological control such as through parasitoids (Spinner *et al.* 2011) or entomopathogenic nematodes (Langford *et al.* 2014). Molecular research has identified microsatellites, mitochondrial genes and ITS sequences for population and phylogenetic studies (Kinnear *et al.* 1998, Morrow *et al.* 2000, Wang *et al.* 2003). Studies have attempted to discern the relationships between *B. tryoni* and other fruit flies for quarantine and biosecurity reasons (Blacket *et al.* 2012, Krosch *et al.* 2012). Of particular interest have been the relationships of *B. tryoni* with sibling species *B. neohumeralis,* including the mating-time dimorphism between the two species (An *et al.* 2002, An *et al.* 2004) and the potential for hybridisation in the field within their sympatric region (Gilchrist and Ling 2006), and with *B. jarvisi*, a more distantly related species that can also hybridise with *B. tryoni* (Cruickshank *et al.* 2001, Shearman *et al.* 2010). Efforts have been made to evaluate SIT parameters such as dispersal (Gilchrist and Meats 2012), effects of sterilising radiation (Collins *et al.* 2008) and mass-reared strain fitness (Meats *et al.* 2004, Weldon 2005). Research into potential improvement to

SIT through the development of a male-only strain have benefitted from identification of genetic and molecular markers, including of sex chromosomal markers (Bennett and Frommer 1997, Shearman *et al.* 2010, Zhao *et al.* 2003a, Zhao *et al.* 2003b), identification of *dsx* and of the sex-determination pathway (Shearman and Frommer 1998), chromosome translocation and cytological studies (Meats *et al.* 2002, Zhao *et al.* 1998) and *B. tryoni* germ line transformation (Raphael *et al.* 2011). A significant advance in the research potential of *Bactrocera* species is imminent, with annotated genomes for *B. tryoni*, *B. neohumeralis* and *B. jarvisi* due to be released (Gilchrist et al., unpublished).

The objective of this study was to use molecular approaches to substantially increase our baseline knowledge in three integrated lines of endeavour, towards better outcomes for pest management strategies that are both specific to *B. tryoni*, but also broadly applicable to other related *Bactrocera* species.

- An investigation of the incidence and prevalence of *Wolbachia* in *B. tryoni* and other Australian tephritid fruit flies combines an important ecological study of this widespread insect bacterium, not previously surveyed in Australian tephritids, with a search for naturally occurring *Wolbachia* strains that may be candidate isolates for future IIT-based approaches.
- 454 pyrosequencing is used for the first time to generate microbiome analyses of Australian tephritid species with different levels of host specialisation, from laboratory lines and field collections. This method benefits from detection of low abundance or rare taxa through deep sequencing to provide data for follow-up functional studies of the role and importance of the microbiome and its application to host fitness and paratransgenic RNAi delivery systems.
- Sequence and expression analysis of sex-determination genes and the genes transcribed in early embryogenesis are required to unravel the sex-determination splicing pathway that starts with the *Dominant Male Determiner* and to identify early zygotic transcripts as potential targets for the generation of a male-only line.

## 7.2 Key findings and future applications

### 7.2.1 *Wolbachia*-based pest control

In order to broaden the range of strategies available for pest management of *B. tryoni*, we considered IIT to be a feasible prospect for investigation. The phenomenon of *Wolbachia*-induced cytoplasmic incompatibility (CI) has been well studied in numerous insects, including another tephritid, *Rhagoletis cerasi* (Arthofer *et al.* 2009b, Boller *et al.* 1976, Riegler and Stauffer 2002). *Wolbachia* strains inducing CI in other species such as *w*Ri in *Drosophila simulans* are present in Australia (Kriesner *et al.* 2013) and appear to also infect other Australian host species (Chapter 3). Significantly, the potential to artificially transfer *Wolbachia* into previously uninfected tephritid species has been demonstrated for *C. capitata* (Zabalou *et al.* 2004) and *B. oleae* (Apostolaki *et al.* 2011); the principle of IIT has been demonstrated in cage experiments for *C. capitata* (Zabalou *et al.* 2004, Zabalou *et al.* 2009); and the fitness of laboratory and GSS *C. capitata* lines carrying the *R. cerasi Wolbachia* strains has been evaluated (Sarakatsanou *et al.* 2011).

IIT requires the release of a male cohort that carries one or more *Wolbachia* strains; the wild population must be either uninfected or infected with a different *Wolbachia* strain that cannot rescue the CI induced by the *Wolbachia* in the released males. To date, no research has surveyed the incidence and prevalence of *Wolbachia* in Australian tephritid fruit flies, an essential first step in assessing the viability of this technique.

*Tropical hotspot for* Wolbachia *horizontal transmission*

PCR screening of 24 species of fruit fly primarily from eastern Australia determined that *Wolbachia* strains were present in eight of those species (Chapter 2). Remarkably, the incidence of *Wolbachia* infection in eight out of 24 species was restricted to northern tropical and subtropical Australia, even within species that were also sampled from their southerly subtropical and temperate distributions. The prevalence of *Wolbachia* infection was fixed in *B. perkinsi*, but low in the other seven species, including the widespread *B. tryoni* and *B. neohumeralis*. These seven fruit fly species carried one or both of two novel *Wolbachia* strains, ST-285 and ST-

289. This extraordinary level of sharing of identical *Wolbachia* strains within a fruit fly community was confirmed by sequence analysis of multiple loci (MLST) and quantification-based allele assignation using allele-specific qPCR (Chapter 3). Furthermore, the fruit fly parasitoid, *Fopius arisanus*, shared *Wolbachia* with two fruit fly individuals: *B. frauenfeldi* and *B. cacuminata*. *Fopius arisanus* is 100% infected by two strains of *Wolbachia*; one appeared to be identical to the CI-causing *w*Ri strain (ST-17); the other identical to a parthenogenesis-causing strain (ST-370) from the parasitoid wasp *Asobara japonica* (Kraaijeveld *et al.* 2011).

These findings clearly identify the Australian tropical fruit fly community as a hot-spot and new model system for the investigation of different routes of horizontal transmission of *Wolbachia* within and between ecological niches. This system has several advantageous characteristics: multiple *Wolbachia* strains that may be transmitted together or independently; potentially different (but currently unknown) reproductive phenotypes; sharing of *Wolbachia* in fruit fly and parasitoid species over two trophic levels; fruit fly species with different host preferences, bioclimatic potential, mating behaviour and ability to hybridise; and native versus introduced species (*F. arisanus* was introduced into Australia in 1956, *B. frauenfeldi* was introduced in 1974) restricting the time-frame of species interactions.

However, as only single fruit fly individuals were found carrying the same combination of *Wolbachia* strains as *F. arisanus*, it is possible that these were PCR-detected as carry-over *Wolbachia* from the parasitoid contact with the fly embryos while ovipositing or probing for oviposition sites (while laboratory contamination has been excluded – Chapters 2 and 3). Furthermore, we did not find *Wolbachia* in any laboratory strains of *B. tryoni*, and were unable to establish infections in laboratory colonies during the time course of this research. Therefore, there remains a possibility that the low levels of infection observed in the tropical fruit fly community might result from continuous exposure of these species to the *Wolbachia* in their environment, creating short-term somatic infections, without germ-line inheritance of the symbiont to this time point.

*Wolbachia – latitudinal gradient*

Tephritid fruit flies screened for *Wolbachia* in this study were sampled from different locations along an extensive latitudinal gradient of eastern Australia. However, *Wolbachia* strains were only detected in northern tropical Australia. There is a species distribution bias to the tropical regions, but widespread species, including *B. tryoni* and *B. neohumeralis*, were *Wolbachia* negative in the more southern temperate regions while they were positive in the northern regions (Chapter 2). The overall low prevalence of *Wolbachia* in *B. tryoni* and *B. neohumeralis* and its restriction to northern Australia may present opportunity for *Wolbachia*-based pest management to be applied in the fruit growing regions of Australia.

One hypothesis for the limited distribution of *Wolbachia* in these tephritid hosts is that *Wolbachia* is responding to the different environmental conditions of the host's latitudinal range. Exposure to high temperatures can reduce *Wolbachia* density within individuals (Bordenstein and Bordenstein 2011) and *Wolbachia* may be unable to proliferate or be efficiently transmitted under such physiological conditions; dryer conditions and cooler winters in temperate climates may result in diminished host fitness (Maes *et al.* 2012, McMeniman and O'Neill 2010); alternatively reduced larval fitness may limit the spread of *Wolbachia* in host populations (Crain *et al.* 2011). Although latitudinal clines in *Wolbachia* prevalence have been found before for *D. melanogaster* (Hoffmann *et al.* 1998) and climatic differences in multiple *Wolbachia* strains in one leaf beetle species (Keller *et al.* 2004); this is the first study to suggest a latitudinal *Wolbachia* distribution in an insect host community rather than in populations of a single species.

An alternative theory for the restriction of *Wolbachia* to the tropics is that we are witnessing a relatively new symbiosis currently invading the population with frequent horizontal transmission with (or without) CI. Turelli and Hoffmann (1991) reported a CI-causing strain (*w*Ri) spreading through Californian populations of *D. simulans* from the south to the north over a three year period and the same *w*Ri strain displaced a non-CI-causing strain (*w*Au) in Australian populations of *D. simulans* along eastern Australia over a 20 year period (Kriesner *et al.* 2013). This precedent predicts that, if these novel strains in Australian tephritids cause CI, the *Wolbachia* may become prevalent in the current locations and spread south to the

extent of the flies' endemic distributions, thus impacting the choice of *Wolbachia* strains used for the implementation of IIT-based technologies.

*CI-causing Wolbachia*

We were unable to establish an infected line of Australian tephritids, and in particular of *B. tryoni* or *B. neohumeralis*; as such we currently have no data to establish whether *Wolbachia* induces CI in these Australian tephritids. The low prevalence of the two native strains of *Wolbachia* (ST-285 and ST-289; Chapter 3) plus the lack of evidence of changes in prevalence over time (Chapter 2) does not indicate a CI phenotype that would necessarily result in a *Wolbachia* sweep through natural populations. The lack of any *Wolbachia* infections in laboratory lines of *B. neohumeralis* originating from tropical north Australia (Appendix C: Table C.1) may also be due to the absence of a CI phenotype, as founder effects and isolation can favour the rapid increase in prevalence of CI inducing strains (Reuter *et al.* 2008). However, these laboratory lines were perhaps not established ideally, i.e. as isofemale lines that would allow low prevalence infections to increase in prevalence; this may be particularly relevant if *Wolbachia* had negative fitness effects on hosts besides the CI phenotype, or induced male-killing. Ongoing screening and selection for *Wolbachia* in newly-founded laboratory lines may be necessary to establish high prevalence lines where reproductive phenotype and transmission rates may be tested. Determination of these phenotypic characteristics is essential for any evaluation of *Wolbachia* candidate strains for pest management. Discovering Australian native and CI-causing *Wolbachia* strains of *Bactrocera* spp. would also present an opportunity to introgress the local *Wolbachia* strains into closely related species or to microinject into a target species if hybridisation is impossible.

*Microinjection*

An alternative to utilising and introgressing *Bactrocera*-native strains of *Wolbachia* involves embryonic microinjection of CI-causing *Wolbachia* from other species such as *Drosophila* spp. or other tephritid fruit flies including *R. cerasi, C. capitata* and *B. oleae* (Apostolaki *et al.* 2011, Riegler and Stauffer 2002, Zabalou *et al.* 2004).

Microinjection of *Wolbachia* was attempted within the time-frame of this study, but stable lines of *B. tryoni* carrying *Wolbachia* strains were not obtained. In this case, a healthy laboratory line of *B. tryoni* (normal egg hatch ~87%) was injected with *Wolbachia* from six sources: *w*MelPop in *D. melanogaster* line *w*1118; *w*Ri in *D. simulans* line DSR; *w*Cer2, native to *R. cerasi*, which had been artificially transferred to *D. simulans* line RC21 (Riegler *et al.* 2004); and three *D. simulans* lines carrying *w*Cer1 and *w*Cer5 (line 3A), and lines 6A and 9D both carrying *w*Cer1 and *w*Cer4 (D. Schneider and W. Miller, unpublished). A cumulative total of 18,187 embryos were injected with embryonic cytoplasm from the above six donor lines over three experimental efforts. Survival to hatching (351 larvae, 1.9%) was poor, and resulted in emergence of 71 females which were set up as isofemale lines. *Wolbachia* were not detected in many of these G0 flies (12 of 71) by PCR-based screening (as used in Chapter 2) and was absent from most of the G1 and G2 flies. However, six lines derived from injection with *w*1118, DSR, RC21, 6A (one isofemale line each) and 9A (two isofemale lines) had PCR-detectable *Wolbachia* in 14-57% of G1 and G2 individuals tested. These *Wolbachia* infections were confirmed by DNA sequencing. Testing of G3 and subsequent generations of flies from these lines failed to detect *Wolbachia.* Overall, these results were too inconclusive to be presented as an experimental chapter, but are included here to point out several features. Embryonic survival was low compared to other microinjection efforts in *B. tryoni* (12.3% emergence reported in Raphael *et al.* 2011) however optimising the preparation of embryos through different brands of bleach for dechorionation, oil and double-sided tape, and less harsh treatment with bleach and desiccation assisted in improving survival in later experiments. Six lines transmitting *Wolbachia* were generated, but were lost at the third generation. There may have been a fitness detriment, not appropriately addressed, that undermined the fecundity of the individuals carrying *Wolbachia* among uninfected individuals (Sarakatsanou *et al.* 2011). A salient example is the transfer of *w*MelPop *Wolbachia* to *Aedes aegypti* mosquitoes after first adaptation to a mosquito cell line; but still some infected lines were lost in early generations through low fecundity or low transmission (McMeniman *et al.* 2009). For the *B. tryoni* experiments, the donor lines were either *D. melanogaster* carrying its adapted *w*MelPop, or *D. simulans* lines carrying either adapted *w*Ri or *Wolbachia* artificially transferred from *R. cerasi*. Adaptation to a more distantly related host (i.e. drosophilids) for over 250 host

generations (from establishment in *D. simulans* in 2000 to microinjection into *B. tryoni* in 2010/2011) may have caused difficulties for adaptation of *w*Cer2 to its new tephritid host,although, in contrast to expectations of required host-adaptation, the *w*Ri strain has been previously transferred directly from *D. simulans* to *Aedes albopictus* (Xi *et al.* 2006). If a laboratory line of *B. tryoni* harbouring its native *Wolbachia* was established from field populations in northern Queensland, microinjection of this strain into uninfected *B. tryoni* embryos would be a useful control to determine if adaptation to host was a factor in the unsuccessful microinjection experiment. Successful direct transfer of *w*Cer2 from *R. cerasi* to *C. capitata* (Zabalou *et al.* 2004) and *B. oleae* was achieved (Apostolaki *et al.* 2011), but *R. cerasi* is not permitted in Australia for quarantine reasons; however it may be possible in the future to import *Wolbachia*-infected *C. capitata*. Although the reproductive phenotype is dependent on the unique host-*Wolbachia* interaction (Jaenike 2007, Poinsot *et al.* 1998, Reynolds *et al.* 2003), the *w*Cer2 strain has been transferred from *R. cerasi* to *D. simulans* (Riegler *et al.* 2004), to *C. capitata* (Zabalou et al. 2004), and to *B. oleae* via *C. capitata* (Apostolaki *et al.* 2011), and in all cases has induced strong CI.

*Bidirectional cytoplasmic incompatibility*

Successful transinfection of *B. tryoni* with *Wolbachia* strains not naturally found in Australian fruit flies may be essential, to ensure CI is induced in a released IIT line by bidirectional incompatibility (Xi *et al.* 2006) and for this method to be applied in endemic areas of northern Australia (Bourtzis and Robinson 2006). The inability of two incompatible strains to rescue the CI induced by the other (Perrot-Minnot *et al.* 1996) may be utilised for an IIT release if, by microinjection of a non-native *Wolbachia* strain, bidirectional incompatibility is strongly induced (Xi *et al.* 2006). One of the main risks of the *Wolbachia* CI-based method is the accidental release of females and subsequent CI-driven *Wolbachia* invasion into the field population, leading to collapse of IIT for that *Wolbachia* strain. The risk is particularly high in a unidirectional CI scenario if coupled with an imperfect GSS. This risk may be mitigated but not eliminated if bidirectional incompatibility occurs, and the females of the release cohort are also incompatible with wild males. A potentially more

reliable method may be to combine *Wolbachia*-based IIT with radiation-based sterilisation of any females not eliminated in the sexing process. This ensures that no females are fertile, the reduced amount of radiation required to sterilise the females is not as damaging to the fitness and competitiveness of the males (Arunachalam and Curtis 1985, Bourtzis and Robinson 2006, Brelsfoard *et al.* 2009, Zabalou *et al.* 2009) and in *Aedes polynesiensis*, *Wolbachia*-induced CI efficiency was not diminished after radiation treatment (Brelsfoard *et al.* 2009). At the same time, sterile released flies carrying *Wolbachia* may be easily recognisable with *Wolbachia* as a simple PCR-based diagnostic marker for differentiation of released and field flies.

### 7.2.2 *Wolbachia*-based pest control – Future research

There is a strong need for a variety of methods to control devastating pests, such as *B. tryoni*, to continually optimise access to the most effective (i.e. cost-effective and sustainable) methods. Therefore, new methods such as IIT should be investigated. Based on the work reported in this thesis, there is potential for implementation of IIT targeting Australian fruit flies such as *B. tryoni*. However, further research into the assessment and implementation of this technique is warranted in the following areas, in particular to:

- Investigate the reproductive phenotype and other physiological parameters of *B. tryoni* and *B. neohumeralis* lines harbouring the native *Wolbachia* strains. This will involve field collections of specimens from infected species in northern Australia and selective breeding to establish infected lines, followed by comprehensive cage experiments to evaluate transmission, CI and fitness effects.
- Establish, by microinjection, *B. tryoni* lines carrying characterised *Wolbachia* strains, e.g. from *R. cerasi* (*w*Cer2), *D. melanogaster* (*w*MelPop) and *D. simulans* (*w*Ri), and as above, assess phenotype, transmission and fitness parameters.
- Assess bidirectional compatibility of fly lines harbouring different *Wolbachia* strains.

- Measure, in the event of a suitable CI-causing *Wolbachia* being found, the competitive fitness of flies, compatibility with genetic sexing techniques, effects of radiation on *Wolbachia* titre and CI-induction.
- Monitor the prevalence of *Wolbachia* in fruit fly species in eastern Australia in order to gain temporal information on *Wolbachia* spread and the mechanism underlying the latitudinal gradient and sharing of *Wolbachia* strains between different species.
- Build on our understanding of *Wolbachia* biology and ecology through further examination of the tropical fruit fly community, including by screening other parasitoid species found in northern Australia for *Wolbachia* to investigate the tropical community for *Wolbachia* horizontal transfer.

### 7.2.3 Microbial and paratransgenic applications to Australian fruit flies

*Inconstant gut microbes*

This study sought to analyse the bacteria associated with tephritid fruit fly species using 454 pyrosequencing (Chapter 4). The advantage of this technique is the accurate relative representation of diversity and high coverage, making it possible to detect rare bacteria, although the relatively short reads limit confident taxonomic assignment. Accordance between the literature and the results presented here strongly suggests the overall abundance of few (up to four) highly represented bacterial families in Australian *Bactrocera* and other tephritids. However, it seems that these bacterial families, also identified by other authors as colonisers of the tephritid gut, vary in abundance and taxonomy at the finer scale, and thus may be opportunistic. Our microbiome data suggested that environment (particular laboratory environment, including diet) is a significant driver of microbial composition between closely related species, such as *B. tryoni* and *B. neohumeralis*; for more distantly-related species genetic distance was identified as a greater factor. However, since environment and diet are clearly a factor, it is probable that the microbiome of flies is also constrained by the host plant associations and diets. Therefore, field caught *B. tryoni* from different host fruits may have a different microbial community to those in the samples collected from citrus plants, examined here.

Furthermore, our study showed differences between three *B. tryoni* laboratory strains housed under identical conditions (see Figure 4.3), which matches substantial variation in replicate samples for some other analyses in drosophilids (Staubach *et al.* 2013, Wong *et al.* 2011), so caution is yet required in making conclusions.

Some limitations of this study derive from a partial lack of true biological and technical replication; so far this appears to be a common problem with pyrosequencing studies of insect communities (Kautz *et al.* 2013, Staubach *et al.* 2013, Wong *et al.* 2011, Wong *et al.* 2013), and perhaps signifies the novelty (and more substantial sequencing costs) of the research area of microbial community ecology based on pyrosequencing methods.

While keeping this limitation in mind, this study looked at the composition and abundance of bacteria in major polyphagous pest species *B. tryoni, B. neohumeralis* and *B. jarvisi,* in comparison to monophagous *B. cacuminata* which has a very restricted host plant range, and the ecologically distinct *D. pornia*. These data suggest greater diversity in the polyphagous *Bactrocera* fruit flies, but no consistent difference has yet been found between *B. tryoni* and *B. neohumeralis* to explain the differences between the distribution of these two species that are otherwise very similar genetically and in terms of their host fruit preferences (Fitt and O'Brien 1985).

No functional or attraction experiments were carried out in this study, instead compositional data were produced as a baseline for further research and the development of pest management strategies. Several avenues are available to employ bacteria that attract fruit flies, for example in trapping, or to provision sterilised flies with a pro-biotic diet to improve their fitness upon release, or simply to supplement the loss of microbial diversity observed in laboratory-maintained lines compared to natural populations (Chapter 4; Chandler *et al.* 2011, Staubach *et al.* 2013). Although the lower fitness of mass-reared lines for SIT is predominantly due to loss of genetic diversity during laboratory-adaptation (Gilchrist *et al.* 2012), loss of microbial diversity within these lines may be counteracted, to provide fitness and field-performance benefits.

An exciting application of microbes inhabiting fruit fly species is bacterial expression of gene constructs as delivery mechanisms for characteristics favourable

to fruit fly hosts (= paratransgenesis). Previous studies have investigated the potential for ingested bacteria expressing dsRNA to be used as species-specific insecticides (Tian *et al.* 2009, Whyard *et al.* 2009). Such systems might be honed to deliver RNAi-based sex-conversion to early embryos. The expression of dsRNA for *tra*, *tra-2* or presumably *M*, could be applied to producing a male-only strain if the delivery and expression were to occur in early embryogenesis while sex is being determined. Although no gut bacteria are detected inside the *D. melanogaster* egg (Bakula 1969, Wong *et al.* 2011), we know that *B. oleae* maternally transmits bacteria essential for larval health to the egg as it is laid, by smearing the egg with bacteria, which then enters through the micropyle (Capuzzo *et al.* 2005). Expression of sex-determination transgenes by bacteria, present in the early embryo, may offer one avenue of RNAi delivery to convert phenotypic females to males. Paratransgenesis, in contrast to developing and releasing transgenic flies, would benefit from increased flexibility in the mass-release line used, which could be more easily outbred to maintain higher fitness levels.

### 7.2.4 Future research – microbial and paratransgenic applications to Australian fruit flies

Delivery of desirable characteristics through transgenic bacteria has the advantage of a well-characterised bacterial transformation system, and delivering an RNAi-based knockout system can be designed very specifically to a particular species. Furthermore, the mechanism for inducing sex-conversion at an early stage may be coupled with lethality linked to a different (e.g. heat shock) promoter to eliminate the transgenic bacteria at the larval stage. However, development of such a technique is in its infancy for fruit flies, and may require substantial optimisation for early embryonic delivery. If successful, this technique may assuage concerns over GMO release, as its application could be restricted to the laboratory environment. Future research recommendations include:

- Investigation of the transmission of microbes from the mother to the egg in *B. tryoni*, or from the environment.
- Evaluation of the fitness benefits or costs of microbes to the different ontogenetic host stages.

- Assessment of the behavioural manipulation of fruit flies by their microbiome.
- Characterisation of the microbiome of field populations of *B. tryoni* and *B. neohumeralis* from areas within their sympatric range, to investigate differences that may contribute to the superior adaptability of *B. tryoni* to dryer and cooler conditions.
- Investigation of the apparent differences in microbial diversity in polyphagous and monophagous fruit flies, by sampling more populations and species, including field populations.
- Quantification of the environmentally-determined component of the fruit fly microbiome (e.g. determined by host plant rather than host genetics).
- Metagenomic and metabolomics analysis of the microbiome of tephritid fruit flies (as an improvement to the limited interpretation based on relatively short and conserved 16S rDNA amplicons).
- Carry out risk assessments associated with the application of transgenic bacteria in paratransgenic approaches to pest management.

### 7.2.5 Manipulating genomes for improved tephritid fruit fly control

The development of molecular tools to improve the efficacy of SIT and IIT can be achieved in different ways using transgenic and non-transgenic methods. Extensive non-transgenic mutagenesis and selection efforts in *C. capitata* led to the generation of suitable Y-chromosome translocations and phenotypic markers (Franz *et al.* 1994), that yielded the first male-only strains based on pupal colour (first generation GSS). Subsequently, the fortunate and serendipitous discovery of a suitable embryonic temperature-sensitive lethal mutation, combined with years of optimisation, allowed the generation of male cohorts (second generation GSS; Franz *et al.* 1994). The current *C. capitata* strain for SIT programmes, Vienna 8, has a chromosomal inversion which suppresses recombination; and applying the Filter Rearing System eliminates any rare recombinants resulting in stable mass rearing (Franz 2005). These methods using mutagenesis are specific to the particular species, although mutations could be introgressed if they were developed in a closely related species. A male-only strain for *B. tryoni* was generated by the same procedures, but

was rendered unsuitable by the discovery of an additional lethal locus (Meats *et al.* 2002). By comparison, the utilisation of transgenics – the stable, heritable insertion of genetic constructs into a host chromosome – have the advantage of transferability across different species with minimal or no alterations required to the transgene. The stable insertion of transgenic gene constructs into the host chromosome has been demonstrated in *B. tryoni* (Raphael *et al.* 2011). Gene constructs inducing female lethality (Schetelig and Handler 2012a) or female to male conversion (Saccone *et al.* 2007) may be used in combination with standard sterilisation induced by radiation or *Wolbachia.* The competitiveness of male flies may be reduced by *Wolbachia* infection (Sarakatsanou *et al.* 2011) or following irradiation (Moreno *et al.* 1991, Toledo *et al.* 2004), but it appears the negative impact of mass rearing and handling may be greater than that of the radiation treatment (Collins *et al.* 2009, Weldon 2005).

An alternative is to combine the male-only transgenic technology with transgenic embryonic lethality (Schetelig *et al.* 2007, Schetelig *et al.* 2009) which produces non-viable offspring when males are released to mate with wild females. However, approaches utilising a single lethal gene are susceptible to the acquisition of resistance, whereas random mutations are generated by irradiation (Robinson 2005). The release of insects carrying a dominant lethal (RIDL) method is another alternative that produces male-only offspring for release, and viable males in matings with wild females after release (Thomas *et al.* 2000). However, the viability of these males means strict safeguards for GMOs must be in place. Paratransgenic delivery of RNAi-based sex-conversion is an alternative to germ-line transformation of flies destined for field release. These strategies may be widely applicable after manageable species-specific modifications to a general gene construct, primarily in the promoter sequences for optimal expression (Figure 7.1).

*Application of sex-determination genes to fruit fly control*

This work describes a substantial increase in the body of molecular information attached to two Australian fruit fly species, *B. tryoni* and *B. jarvisi*. Specific information for the primary sex-determination genes and their compatible expression

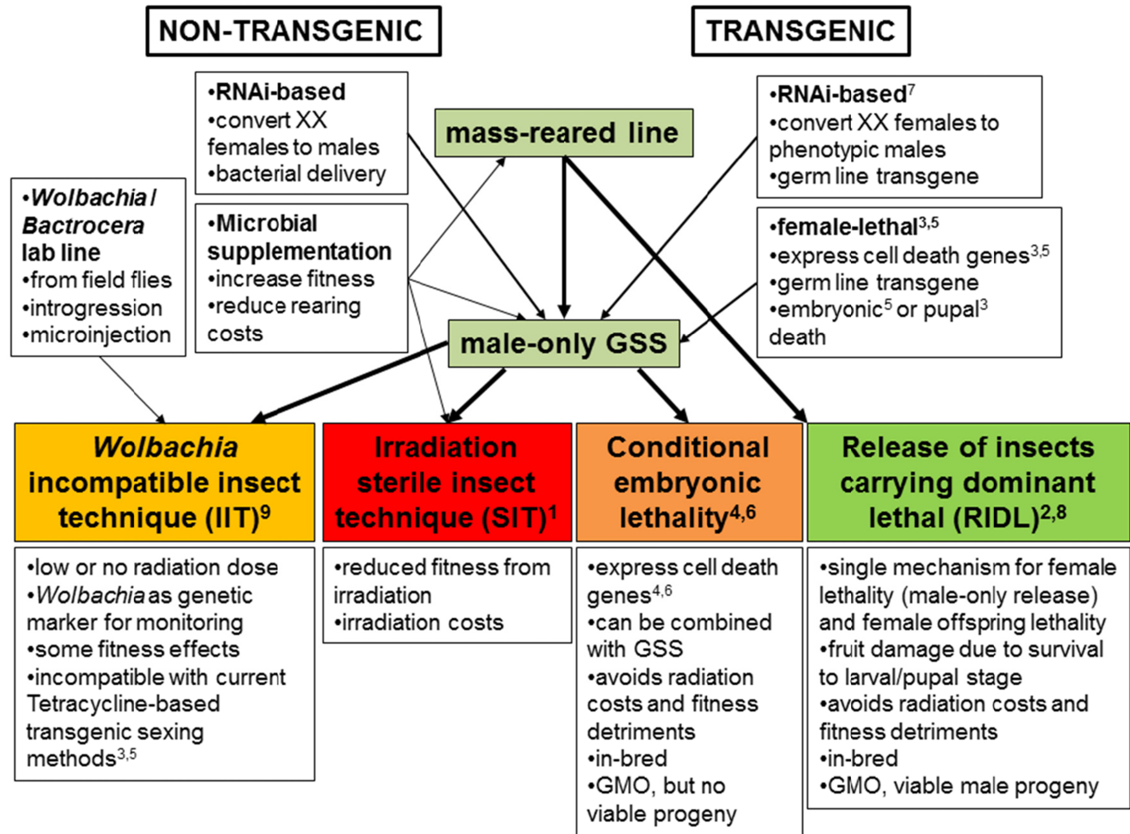



**Figure 7.1** Diagram depicts transgenic and non-transgenic methods for improving the current pest management repertoire, and the associated advantages and disadvantages. (References 1-9: Dyck *et al.* 2005, Fu *et al.* 2007, Heinrich and Scott 2000, Schetelig *et al.* 2007, Schetelig *et al.* 2009, Schetelig and Handler 2012a, Schetelig and Handler 2012b, Schetelig *et al.* 2012, Thomas *et al.* 2000, Zabalou *et al.* 2009)

pattern in both species, and the whole *B. jarvisi* poly(A)$^+$ transcriptome present in young embryos at two consecutive time periods in male and female samples is recorded here for the first time.

Standard molecular methods were used to obtain genomic and mRNA sequence for genes most important to the sex-determination pathway in *B. tryoni* and *B. jarvisi*, namely *Sxl*, *tra* and *tra-2* (Chapter 5). The functionality of female-specific transcripts of *tra* and non-sex-specific *tra-2* in female development has been demonstrated by RNAi in *C. capitata*, *B. oleae* and *A. suspensa* (Lagos *et al.* 2007, Pane *et al.* 2002, Salvemini *et al.* 2009, Schetelig *et al.* 2012). Preliminary assessment of Bttra and Bttra-2 suggests the same functionality – injection of full-length Bttra and Bttra-2

dsRNA into B. tryoni embryos of 3h and 7h AEL was performed, with seven intersexes recovered from Bttra-2 injections into 7h embryos suggesting incomplete phenotypic conversion; and complete conversion resulting from Bttra injections of 3h embryos, whereby two phenotypic males developed that sired all female offspring when mated to a normal female (J.L. Morrow, K. Raphael, M. Frommer and D.C.A. Shearman, unpublished). Therefore, conditional transgenic RNAi or bacterial delivery of RNAi-based strategies for generating male-only strains may be an effective means for improving SIT and IIT. Such an approach would come with the benefit and cost-effectiveness of converting and not killing the females, unlike conditional lethal protocols. In addition, RNAi can be highly species-specific, restricting concerns of GMO releases, and systems where both *tra* and *tra-2* transcripts can be simultaneously knocked-down (Schetelig *et al.* 2012), reduce the potential loss of effectiveness in the event of a mutation in either one of the RNAs.

As described in Chapter 5, quantitative RT-PCR calculated the relative expression of *Sxl*, *tra*, *tra-2* and *dsx* as *B. jarvisi* embryos developed from 1-9h, demonstrating that *Sxl* increased in transcript level in both male and female at 3-4h AEL; and that *tra-2* expression remained essentially constant. Most telling, however, was that in males, $tra^{F}$ and $dsx^{F}$ transcripts decreased in concentration concurrently, indicating that the specifically spliced transcripts were being degraded and new transcripts were not being spliced into the female-specific form. We suggested that, because this occurs simultaneously in the transcripts that rely on the TRA/TRA-2 spliceosome for female-specific splicing, that a single mechanism is inhibiting the spliceosome, enacted by *M* or its proxy. Interestingly, in the *B. tryoni* embryos tested, these time points were delayed, but appeared to follow the same early progression.

The dynamic gene expression during the pre-blastoderm development in embryos justified the decision to use qRT-PCR first on known sex-determination genes to determine an appropriate developmental time window to examine the poly(A)$^{+}$ transcriptome using deep-sequencing. The focus of transcriptome analysis was on the expression of 28 sex-determination genes described in *D. melanogaster*, of which 15 new homologues were identified for *B. jarvisi* (Chapter 6). Two genes were of particular interest due to their expression profiles. Transcripts of *B. jarvisi sisterless A* appeared very early in development and increased in abundance rapidly, and may provide suitable candidate promoter sequences for transgenics. *Bactrocera*

*jarvisi extra-macrochaetae* was zygotically transcribed and displayed a different expression profile to maternally-derived *emc* in *D. melanogaster.* While *emc* was up-regulated in both male and female samples more than 100-fold in 3-5h embryos, the different regulation of this gene when compared to *D. melanogaster* warrants further investigation into its role in early development of tephritid fruit flies.

*Transcriptome assembly and analysis*

The transcriptome data presented in this thesis (Chapter 6) may be mined more comprehensively and effectively in several ways. Both Trinity *de novo* and CLC Genomics Workbench *de novo* generated assemblies with similar numbers of contigs, however the Trinity assembly was not fully analysed due to time constraints. Comparing different methods of assembly was beyond the scope of this thesis, but a logical next step will be to reassess the trimming, assembly and mapping parameters, and the various tools available to achieve optimal outcomes. The CLC assembly did not assemble full ORFs in all cases (Table 6.4) and may be improved by more stringent quality control; however in some cases this discrepancy may be attributed to heterozygosity or gene families confounding the assembly.

The qRT-PCR validated genes (Chapter 5) were essential to the assessment and validation of the quality of the CLC assembly. Two male sample libraries, one at each time period, were problematic because principal co-ordinates analysis found these two samples did not cluster with their replicates and did not emulate the expression profile for validated genes. These two libraries were eliminated from the analyses, thus reducing the power of the differential expression.

Three more options for analysing the transcriptome data are now becoming available, taking advantage of the *B. jarvisi* genome assembly (Gilchrist et al. unpublished). First, mapping the RNA-Seq output directly to the *B. jarvisi* genome; second, genome-guided assembly of the RNA-Seq output; and third mapping the transcriptome directly to Y-chromosome DNA sequences from the annotated genome, derived from subtraction of genomic sequence of females from males. The third method may lead directly to sequences transcribed from the Y-chromosome, but a combination of the first two methods plus the *de novo* assembly has been used

successfully to isolate sequences from the W-chromosome in chickens (Ayers *et al.* 2013). All of these methods may be analysed to ascertain differentially expressed transcripts, but the genome-based assemblies may discern different splicing events in males and females that were not detected using *de novo* assembly.

*Highly up-regulated zygotic transcripts*

No strong candidates for Y-chromosome transcripts were identified, as this would require sequences to be up-regulated in males over time and absent in females. Nevertheless, discovery of transcripts increasing in expression as embryos age has identified potential candidate sequences strongly expressed in the early embryo; these may subsequently be located on the genome along with promoter and enhancer sequences. Literature searches identified *slam*, *sryα*, *bnk* and *nullo* as potential sources of promoter and enhancer elements for transgene construction, as they need to be species-specific and expressed strongly in the early embryo. These elements have been used in *A. suspensa* to drive Tet-off transgenic embryonic sexing system (Schetelig and Handler 2012a). This technique utilises the sex-specific splicing of the *tra* transcript linked to a cell death effector gene, such that male-specific inclusion of introns encoding stop codons prevents production of the lethal product; in females the lethal product is formed. This method in its current form is incompatible with *Wolbachia*-based IIT, as rearing the release line on a diet containing tetracycline will eliminate the *Wolbachia*. Therefore, an alternative to the Tet-transactivator system is required.

In the *B. jarvisi* transcriptome, only *slam*, *nullo* and *sryα* were found, and *sryα* did not exhibit an ideal expression profile. Nevertheless, *slam* and *nullo* are good candidates for transgenic constructs that require an early promoter. Other candidates from the transcriptome presented here include *sisA* and *emc*.

The utility of suitable promoter/enhancer sequences is further demonstrated by the potential of conditional embryonic lethality systems. The *sryα* promoter/enhancer element was used in *C. capitata* and *A. suspensa* to drive a cell death gene for conditional embryonic lethal systems (Schetelig *et al.* 2009, Schetelig and Handler 2012b). For a practical application, transgenic males mating with wild females result

in embryonic death. The advantage of this line is the improved fitness of the flies and reduced cost by avoiding radiation. It is however essential that lethality occurs at the embryonic stage, as death at later stages is not satisfactory when most of the damage is caused by larvae. Transgenes often experience positional effects, affecting the success of the system, and line fitness is an important parameter to be optimised.

### 7.2.6 Future research – manipulating genomes for improved tephritid fruit fly control

The most immediate improvement to SIT is a male-only line, providing cost-benefits to rearing of strains and likely improvements to the effectiveness of released sterile males in the field. The search for the *Dominant Male Determiner* is a key undertaking and pivotal in the understanding of the evolutionary relationships between organisms that share the same mechanisms, as well as those that have modifications. Identification of the *M* transcript (which may be protein coding, non-coding mRNAs or microRNAs) that defines male development provides an ideal candidate for male-only strains.

Furthermore, there may be an effect of *Wolbachia* infection on sex-determination pathway genes that may be explored. Specifically, *Wolbachia* have been implicated in the restoration of females carrying a mutation in *Sex-lethal*, a key gene not only in *D. melanogaster* somatic sex determination, but also dosage compensation and germ line development (Starr and Cline 2002). While *Sex-lethal* does not appear to embrace the same important functions in Tephritidae, studying the effect of *Wolbachia* on sex-determination genes, in particular genes involved in germ-line sex determination, may provide exciting new insights into the regulation and function of these genes as well as evidence for the place of *Wolbachia* in manipulating host expression during early development.

Further research in this area is warranted to:

- Reassemble transcriptome sequences with alternative parameters to optimise assembly and differential expression analysis.

- Perform genome-guided assembly of the *B. jarvisi* transcriptome and mapping the *B. jarvisi* transcriptome directly to the genome, and repeat differential expression analysis.
- Use *B. jarvisi* Y-chromosome genome to map transcriptome and determine if any transcripts are expressed from the Y-chromosome.
- Investigate *BjsisA* and *Bjemc* further by qRT-PCR to validate the expression profiles.
- Characterise zygotic expression further, sequence the non-protein-coding fraction of the genome using the same technique for separating male and female embryos, to identify important transcripts.
- Identify promoter regions of early embryonic genes for transgene construction in *B. jarvisi* and consequently, *B. tryoni*.
- Develop a genetic sexing transgenic construct using *B. tryoni*-specific sequences based on the model applied to *C. capitata* and *A. suspensa*
- Investigate transgenic systems to control expression as an alternative to the tetracycline-based control, to use in *Wolbachia*-carrying lines.
- Investigate the impact of *Wolbachia* on expression of sex-determination genes.

### 7.3 General Conclusions

This thesis has detailed molecular research into three aspects of tephritid pest control. *Wolbachia*-based IIT deserves further investigation as outlined in the previous sections. Microbiome studies begin to address fruit fly fitness as determined by the microbial composition. Sequences for sex-transforming genes in *B. tryoni* and *B. jarvisi* are now available, preliminary RNAi functional studies in *B. tryoni* indicated that the expression of dsRNA of either *tra* or *tra-2* at the pre-blastoderm stage will transform genotypic females into phenotypic males, promoter / enhancer elements can be obtained from the genomes of *B. tryoni*, *B. neohumeralis* and *B. jarvisi* by homology to the transcriptome sequences of early zygotically expressed genes in *B. jarvisi*. This research is neatly positioned in the quest for improved pest management strategies for tephritid fruit flies.

# References

Ahantarig A, Trinachartvanit W, Kittayapong P (2008). Relative *Wolbachia* density of field-collected *Aedes albopictus* mosquitoes in Thailand. *J Vect Ecol* **33:** 173-177.

Aharon Y, Pasternak Z, Ben-Yosef M, Behar A, Lauzon CR, Yuval B *et al* (2013). Phylogenetic, metabolic, and taxonomic diversities shape mediterranean fruit fly microbiotas during ontogeny. *Appl Environ Microbiol* **79:** 303-313.

Ahmed MZ, Greyvenstein OFC, Erasmus C, Welch JJ, Greeff JM (2013). Consistently high incidence of *Wolbachia* in global fig wasp communities. *Ecol Entomol* **38:** 147-154.

Amrein H, Gorman M, Nöthiger R (1988). The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell* **55:** 1025-1035.

An W, Cho S, Ishii H, Wensink PC (1996). Sex-specific and non-sex-specific oligomerization domains in both of the *doublesex* transcription factors from *Drosophila melanogaster*. *Mol Cell Biol* **16:** 3106-3111.

An X, Wilkes K, Bastian Y, Morrow JL, Frommer M, Raphael KA (2002). The *period* gene in two species of tephritid fruit fly differentiated by mating behaviour. *Insect Mol Biol* **11:** 419-430.

An X, Tebo M, Song S, Frommer M, Raphael KA (2004). The *cryptochrome* (*cry*) gene and a mating isolation mechanism in tephritid fruit flies. *Genetics* **168:** 2025-2036.

Andersen SB, Hansen LH, Sapountzis P, Sørensen SJ, Boomsma JJ (2013). Specificity and stability of the *Acromyrmex–Pseudonocardia* symbiosis. *Mol Ecol* **22:** 4307-4321.

Anderson DT (1962). The Embryology of *Dacus tryoni* (Frogg.)[Diptera, Trypetidae (= Tephritidae)], the Queensland Fruit-Fly. *Development* **10:** 248-292.

Andrewartha HG, Monro J, Richardson NL (1967). The use of sterile males to control populations of Queensland fruit fly *Dacus tryoni* (Frogg.)(Diptera: Tephritidae) II. Field experiments in New South Wales. *Aust J Zool* **15:** 475-499.

Apostolaki A, Livadaras I, Saridaki A, Chrysargyris A, Savakis C, Bourtzis K (2011). Transinfection of the olive fruit fly *Bactrocera oleae* with *Wolbachia*: towards a symbiont-based population control strategy. *J App Ent* **135:** 546-553.

Arakaki N, Miyoshi T, Noda H (2001). *Wolbachia*-mediated parthenogenesis in the predatory thrips *Franklinothrips vespiformis* (Thysanoptera: Insecta). *Proc R Soc Lond B Biol Sci* **268:** 1011-1016.

Arbeitman MN, Furlong EEM, Imam F, Johnson E, Null BH, Baker BS *et al* (2002). Gene expression during the life cycle of *Drosophila melanogaster*. *Science* **297:** 2270-2275.

Arthofer W, Riegler M, Avtzis DN, Stauffer C (2009a). Evidence for low-titre infections in insect symbiosis: *Wolbachia* in the bark beetle *Pityogenes chalcographus* (Coleoptera, Scolytinae). *Environ Microbiol* **11:** 1923-1933.

Arthofer W, Riegler M, Schneider D, Krammer M, Miller WJ, Stauffer C (2009b). Hidden *Wolbachia* diversity in field populations of the European cherry fruit fly, *Rhagoletis cerasi* (Diptera, Tephritidae). *Mol Ecol* **18:** 3816-3830.

Arthofer W, Riegler M, Schuler H, Schneider D, Moder K, Miller WJ *et al* (2011). Allele intersection analysis: A novel tool for multi locus sequence assignment in multiply infected hosts. *PLoS One* **6:** e22198.

Arunachalam N, Curtis CF (1985). Integration of radiation with cytoplasmic incompatibility for genetic control in the *Culex pipiens* complex (Diptera: Culicidae). *J Med Entomol* **22:** 648-653.

Austin AD, Yeates DK, Cassis G, Fletcher MJ, La Salle J, Lawrence JF *et al* (2004). Insects 'Down Under'–Diversity, endemism and evolution of the Australian insect fauna: examples from select orders. *Aust J Entomol* **43:** 216-234.

Avila FW, Erickson JW (2007). *Drosophila* JAK/STAT Pathway Reveals Distinct Initiation and Reinforcement Steps in Early Transcription of *Sxl*. *Curr Biol* **17:** 643-648.

Ayers KL, Davidson NM, Demiyah D, Roeszler KN, Grutzner F, Sinclair AH *et al* (2013). RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken embryos and allows comprehensive annotation of W-chromosome genes. *Genome Biology* **14:** R26.

Bakula M (1969). The persistence of a microbial flora during postembryogenesis of *Drosophila melanogaster*. *J Invertebr Pathol* **14:** 365-374.

Baldo L, Bordenstein S, Wernegreen JJ, Werren JH (2006a). Widespread recombination throughout *Wolbachia* genomes. *Mol Biol Evol* **23:** 437-449.

Baldo L, Dunning Hotopp JC, Jolley KA, Bordenstein SR, Biber SA, Choudhury RR *et al* (2006b). Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Appl Environ Microbiol* **72:** 7098-7110.

Baldo L, Werren JH (2007). Revisiting *Wolbachia* supergroup typing based on WSP: Spurious lineages and discordance with MLST. *Curr Microbiol* **55:** 81-87.

Baldo L, Ayoub NA, Hayashi CY, Russell JA, Stahlhut JK, Werren JH (2008). Insight into the routes of *Wolbachia* invasion: high levels of horizontal transfer in the spider genus *Agelenopsis* revealed by *Wolbachia* strain and mitochondrial DNA diversity. *Mol Ecol* **17:** 557-569.

Bandziulis RJ, Swanson MS, Dreyfuss G (1989). RNA-binding proteins as developmental regulators. *Genes Dev* **3:** 431-437.

Barbash DA, Cline TW (1995). Genetic and molecular analysis of the autosomal component of the primary sex determination signal of *Drosophila melanogaster*. *Genetics* **141:** 1451-1471.

Bateman MA (1972). The ecology of fruit flies. *Annu Rev Entomol* **17:** 493-518.

Beard CB, O'Neill SL, Tesh RB, Richards FF, Aksoy S (1993). Modification of arthropod vector competence via symbiotic bacteria. *Parasitol Today* **9:** 179-183.

Behar A, Yuval B, Jurkevitch E (2005). Enterobacteria-mediated nitrogen fixation in natural populations of the fruit fly *Ceratitis capitata*. *Mol Ecol* **14:** 2637-2643.

Behar A, Jurkevitch E, Yuval B (2008a). Bringing back the fruit into fruit fly–bacteria interactions. *Mol Ecol* **17:** 1375-1386.

Behar A, Yuval B, Jurkevitch E (2008b). Gut bacterial communities in the Mediterranean fruit fly *Ceratitis capitata* and their impact on host longevity. *J Insect Physiol* **54:** 1377-1383.

Behar A, Ben-Yosef M, Lauzon CR, Yuval B, Jurkevich E (2009). Structure and function of the bacterial community associated with the Mediterranean fruit fly. In: Bourtzis K, Miller TA (eds). *Insect Symbiosis*. pp 251-271.

Belote JM, Baker BS (1982). Sex determination in *Drosophila melanogaster*: analysis of *transformer-2*, a sex-transforming locus. *Proc Natl Acad Sci USA* **79:** 1568-1572.

Ben-Yosef M, Aharon Y, Jurkevitch E, Yuval B (2010). Give us the tools and we will do the job: symbiotic bacteria affect olive fly fitness in a diet-dependent fashion. *Proc R Soc Lond B Biol Sci* **277:** 1545-1552.

Ben Ami E, Yuval B, Jurkevitch E (2010). Manipulation of the microbiota of mass-reared Mediterranean fruit flies *Ceratitis capitata* (Diptera: Tephritidae) improves sterile male sexual performance. *ISME J* **4:** 28-37.

Bennett CL, Frommer M (1997). The *white* gene of the tephritid fruit fly *Bactrocera tryoni* is characterized by a long untranslated 5' leader and a 12kb first intron. *Insect Mol Biol* **6:** 343-356.

Blacket MJ, Semeraro L, Malipatil MB (2012). Barcoding Queensland fruit flies (*Bactrocera tryoni*): Impediments and improvements. *Mol Ecol Res* **12:** 428-436.

Blümel S, Russ K (1989). Manipulation of races. *Fruit Flies: Their Biology, Natural Enemies and Control*. Elsevier. pp 387–389.

Boggs RT, Gregor P, Idriss S, Belote JM, McKeown M (1987). Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the *transformer* gene. *Cell* **50:** 739-747.

Boller EF, Russ K, Vallo V, Bush GL (1976). Incompatible races of European cherry fruit fly, *Rhagoletis cerasi* (Diptera: Tephritidae), their origin and potential use in biological control. *Entomol Exp Appl* **20:** 237-247.

Bordenstein SR, O'Hara FP, Werren JH (2001). *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature* **409:** 707-710.

Bordenstein SR, Bordenstein SR (2011). Temperature affects the tripartite interactions between bacteriophage WO, *Wolbachia*, and cytoplasmic incompatibility. *PLoS One* **6:** e29106.

Bourtzis K, Dobson SL, Braig HR, O'Neill SL (1998). Rescuing *Wolbachia* have been overlooked. *Nature* **391:** 852-853.

Bourtzis K, Robinson AS (2006). Insect pest control using *Wolbachia* and/or radiation. In: Bourtzis K, Miller TA (eds). *Insect Symbiosis*. CRC Press, Taylor and Francis group: Florida, USA. pp 225-246.

Bousch MG, Matsumura F (1967). Insecticidal Degradation by *Pseudomonas melophthora*, the Bacterial Symbiote of the Apple Maggot. *J Econ Entomol* **60:** 918-920.

Bownes M (1975). A photographic study of development in the living embryo of *Drosophila melanogaster*. *J Embryol Exp Morphol* **33:** 789-801.

Boyle L, O'Neill SL, Robertson HM, Karr TL (1993). Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*. *Science* **260:** 1796-1799.

Braig HR, Guzman H, Tesh RB, O'Neill SL (1994). Replacement of the natural *Wolbachia* symbiont of *Drosophila simulans* with a mosquito counterpart. *Nature* **367:** 453-455.

Braig HR, Zhou W, Dobson SL, O'Neill SL (1998). Cloning and characterization of a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia pipientis*. *J Bacteriol* **180:** 2373-2378.

Breeuwer JAJ, Werren JH (1993). Cytoplasmic incompatibility and bacterial density in *Nasonia vitripennis*. *Genetics* **135:** 565-574.

Brelsfoard CL, St Clair W, Dobson SL (2009). Integration of irradiation with cytoplasmic incompatibility to facilitate a lymphatic filariasis vector elimination approach. *Parasit Vectors* **2:** 38.

Broderick NA, Raffa KF, Handelsman J (2006). Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proc Natl Acad Sci USA* **103:** 15196-15199.

Broderick NA, Lemaitre B (2012). Gut-associated microbes of *Drosophila melanogaster*. *Gut microbes* **3:** 307-321.

Brownlie JC, Cass BN, Riegler M, Witsenburg JJ, Iturbe-Ormaetxe I, McGraw EA *et al* (2009). Evidence for metabolic provisioning by a common invertebrate endosymbiont, *Wolbachia pipientis*, during periods of nutritional stress. *PLoS Pathog* **5:** e1000368.

Burghardt G, Hediger M, Siegenthaler C, Moser M, Dubendorfer A, Bopp D (2005). The *transformer2* gene in *Musca domestica* is required for selecting and maintaining the female pathway of development. *Dev Genes Evol* **215:** 165-176.

Burtis KC, Baker BS (1989). *Drosophila doublesex* gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* **56:** 997-1010.

Burtis KC, Coschigano KT, Baker BS, Wensink PC (1991). The *doublesex* proteins of *Drosophila melanogaster* bind directly to a sex-specific yolk protein gene enhancer. *EMBO J* **10:** 2577-2582.

Bushati N, Stark A, Brennecke J, Cohen SM (2008). Temporal Reciprocity of miRNAs and Their Targets during the Maternal-to-Zygotic Transition in *Drosophila*. *Curr Biol* **18:** 501-506.

Cameron EC, Sved JA, Gilchrist AS (2010). Pest fruit fly (Diptera: Tephritidae) in northwestern Australia: One species or two? *Bull Entomol Res* **100:** 197-206.

Campuzano S (2001). *Emc*, a negative HLH regulator with multiple functions in *Drosophila* development. *Oncogene* **20:** 8299-8307.

Capuzzo C, Firrao G, Mazzon L, Squartini A, Girolami V (2005). '*Candidatus* Erwinia dacicola', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). *Int J Syst Evol Microbiol* **55:** 1641-1647.

Carmichael AE, Wharton RA, Clarke AR (2005). Opiine parasitoids (Hymenoptera: Braconidae) of tropical fruit flies (Diptera: Tephritidae) of the Australian and South Pacific region. *Bull Entomol Res* **95:** 545-569.

Casiraghi M, Anderson TJC, Bandi C, Bazzocchi C, Genchi C (2001). A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of *Wolbachia* endosymbionts. *Parasitology* **122:** 93-103.

Caspi-Fluger A, Inbar M, Mozes-Daube N, Katzir N, Portnoy V, Belausov E *et al* (2012). Horizontal transmission of the insect symbiont *Rickettsia* is plant-mediated. *Proc R Soc Lond B Biol Sci* **279:** 1791-1796.

Ceuppens S, Boon N, Uyttendaele M (2013). Diversity of *Bacillus cereus* group strains is reflected in their broad range of pathogenicity and diverse ecological lifestyles. *FEMS Microbiol Ecol* **84:** 433-450.

Chandler JA, Lang JM, Bhatnagar S, Eisen JA, Kopp A (2011). Bacterial communities of diverse *Drosophila* species: ecological context of a host–microbe model system. *PLoS Genet* **7:** e1002272.

Chao A, Chazdon RL, Colwell RK, Shen TJ (2006). Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* **62:** 361-371.

Charlat S, Riegler M, Baures I, Poinsot D, Stauffer C, Mercot H (2004). Incipient evolution of *Wolbachia* compatibility types. *Evolution* **58:** 1901-1908.

Cho S, Wensink PC (1997). DNA binding by the male and female *doublesex* proteins of *Drosophila melanogaster*. *J Biol Chem* **272:** 3185-3189.

Chouaia B, Rossi P, Montagna M, Ricci I, Crotti E, Damiani C *et al* (2010). Molecular evidence for multiple infections as revealed by typing of *Asaia* bacterial symbionts of four mosquito species. *Appl Environ Microbiol* **76:** 7444-7450.

Christenson LD, Foote RH (1960). Biology of fruit flies. *Annu Rev Entomol* **5:** 171-192.

Clarke AR, Powell KS, Weldon CW, Taylor PW (2011). The ecology of *Bactrocera tryoni* (Diptera: Tephritidae): what do we know to assist pest management? *Ann Appl Biol* **158:** 26-54.

Cline TW (1978). Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with *daughterless*. *Genetics* **90:** 683-698.

Clough E, Oliver B (2012). Genomics of sex determination in *Drosophila*. *Briefings Funct Genomics* **11:** 387-394.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucl Acids Res* **37:** D141-D145.

Collins SR, Weldon CW, Banos C, Taylor PW (2008). Effects of irradiation dose rate on quality and sterility of Queensland fruit flies, *Bactrocera tryoni* (Froggatt). *J App Ent* **132:** 398-405.

Collins SR, Weldon CW, Banos C, Taylor PW (2009). Optimizing irradiation dose for sterility induction and quality of *Bactrocera tryoni*. *J Econ Entomol* **102:** 1791-1800.

Concha C, Scott MJ (2009). Sexual development in *Lucilia cuprina* (Diptera, Calliphoridae) is controlled by the *transformer* gene. *Genetics* **182:** 785-798.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21:** 3674-3676.

Cook C (2005). *The role of cryptochrome in mating time difference in tephritid fruit flies, Bactrocera tryoni and Bactrocera neohumeralis*. University of Sydney: Sydney.

Corby-Harris V, Pontaroli AC, Shimkets LJ, Bennetzen JL, Habel KE, Promislow DEL (2007). Geographical distribution and diversity of bacteria associated with natural populations of *Drosophila melanogaster*. *Appl Environ Microbiol* **73:** 3470-3479.

Coscrato VE, Braz ASK, Perondini ALP, Selivon D, Marino CL (2009). *Wolbachia* in *Anastrepha* fruit flies (Diptera: Tephritidae). *Curr Microbiol* **59:** 295-301.

Cox CR, Gilmore MS (2007). Native microbial colonization of *Drosophila melanogaster* and its use as a model of *Enterococcus faecalis* pathogenesis. *Infection and Immunity* **75:** 1565-1576.

Crain PR, Mains JW, Suh E, Huang Y, Crowley PH, Dobson SL (2011). *Wolbachia* infections that reduce immature insect survival: Predicted impacts on population replacement. *BMC Evol Biol* **11:** 290.

Cronmiller C, Schedl P, Cline TW (1988). Molecular characterization of *daughterless*, a *Drosophila* sex determination gene with multiple roles in development. *Genes Dev* **2:** 1666-1676.

Cruickshank L, Jessup AJ, Cruickshank DJ (2001). Interspecific crosses of *Bactrocera tryoni* (Froggatt) and *Bactrocera jarvisi* (Tryon) (Diptera:Tephritidae) in the laboratory. *Aust J Entomol* **40:** 278-280.

Curthoys S (1997). Investigating phylogenetic relationships within the dacine fruit flies using DNA sequence comparisons. Bachelor of Science (Honours) thesis, University of Sydney, Sydney.

De Renzis S, Elemento O, Tavazoie S, Wieschaus EF (2007). Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. *PLoS Biol* **5:** e117.

Dobson SL, Bourtzis K, Braig HR, Jones BF, Zhou W, Rousset F *et al* (1999). *Wolbachia* infections are distributed throughout insect somatic and germ line tissues. *Insect Biochem Mol Biol* **29:** 153-160.

Dominiak BC, Daniels D (2012). Review of the past and present distribution of Mediterranean fruit fly (*Ceratitis capitata* Wiedemann) and Queensland fruit fly (*Bactrocera tryoni* Froggatt) in Australia. *Aust J Entomol* **51:** 104-115.

Dominiak BC, Ekman JH (2013). The rise and demise of control options for fruit fly in Australia. *Crop Protection* **51:** 57-67.

Doudoumis V, Tsiamis G, Wamwiri F, Brelsfoard C, Alam U, Aksoy E *et al* (2012). Detection and characterization of *Wolbachia* infections in laboratory and natural populations of different species of tsetse flies (genus *Glossina*). *BMC Microbiol* **12:** S3.

Drew RAI, Hooper GHS, Bateman MA (1978). *Economic Fruit Flies of the South Pacific Region*. Entomology Branch, Queensland Department of Primary Industries: Brisbane.

Drew RAI, Hooper GHS (1981). The responses of fruit fly species (Diptera: Tephritidae) in Australia to various attractants. *Aust J Entomol* **20:** 201-205.

Drew RAI, Courtice AC, Teakle DS (1983). Bacteria as a natural source of food for adult fruit flies (Diptera: Tephritidae). *Oecologia* **60:** 279-284.

Drew RAI, Lloyd AC (1987). Relationship of fruit flies (Diptera: Tephritidae) and their bacteria to host plants. *Ann Ent Soc Am* **80:** 629-636.

Drew RAI (1989). The tropical fruit flies (Diptera: Tephritidae: Dacinae) of the Australasian and Oceanian regions. *Memoirs - Queensland Museum* **26:** 1-521.

Dumler JS, Barbet AF, Bekker C, Dasch GA, Palmer GH, Ray SC *et al* (2001). Reorganization of genera in the families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and'HGE agent'as subjective synonyms of *Ehrlichia phagocytophila*. *Int J Syst Evol Microbiol* **51:** 2145-2165.

Dunbar HE, Wilson ACC, Ferguson NR, Moran NA (2007). Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol* **5:** e96.

Dunning Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, Munoz Torres MC *et al* (2007). Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* **317:** 1753-1756.

Duplouy A, Vermenot C, Davies N, Roderick G, Hurst GDD, Charlat S (2009). Assessing risks of *Wolbachia* DNA cross-specimen contamination following mass collection and ethanol storage. *Mol Ecol Res* **9:** 46-50.

Duron O, Bouchon D, Boutin S, Bellamy L, Zhou L, Engelstädter J *et al* (2008). The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biology* **6:** 27.

Dyck VA, Hendrichs J, Robinson AS (eds) (2005) *Sterile insect technique: Principles and practice in area-wide integrated pest management*. Springer: The Netherlands.

Engel P, Moran NA (2013). The gut microbiota of insects - diversity in structure and function. *FEMS Microbiology Reviews* **37:** 699-735.

Erickson JW, Cline TW (1993). A bZIP protein, *sisterless*-a, collaborates with bHLH transcription factors early in *Drosophila* development to determine sex. *Genes Dev* **7:** 1688-1702.

Erickson JW, Quintero JJ (2007). Indirect effects of ploidy suggest X chromosome dose, not the X: A ratio, signals sex in *Drosophila*. *PLoS Biol* **5:** e332.

Estes AM, Hearn DJ, Bronstein JL, Pierson EA (2009). The olive fly endosymbiont, "*Candidatus* Erwinia dacicola", switches from an intracellular existence to an extracellular existence during host insect development. *Appl Environ Microbiol* **75:** 7097-7106.

Fast EM, Toomey ME, Panaram K, Desjardins D, Kolaczyk ED, Frydman HM (2011). *Wolbachia* enhance *Drosophila* stem cell proliferation and target the germline stem cell niche. *Science* **334:** 990-992.

Fay HAC (2012). A highly effective and selective male lure for *Bactrocera jarvisi* (Tryon)(Diptera: Tephritidae). *Aust J Entomol* **51:** 189-197.

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM *et al* (2013). FunGene: the functional gene pipeline and repository. *Front Microbiol* **4:** 291.

Fitt GP, O'Brien RW (1985). Bacteria associated with four species of *Dacus* (Diptera: Tephritidae) and their role in the nutrition of the larvae. *Oecologia* **67:** 447-454.

Fitt GP (1990). Comparative fecundity, clutch size, ovariole number and egg size of *Dacus tryoni* and *D. jarvisi*, and their relationship to body size. *Entomol Exp Appl* **55:** 11-21.

Fletcher B (1987). The biology of dacine fruit flies. *Annu Rev Entomol* **32:** 115-144.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994). DNA primers for amplification of mitochondrial *cytochrome c oxidase* subunit 1 from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **3:** 294-299.

Frago E, Dicke M, Godfray HCJ (2012). Insect symbionts as hidden players in insect–plant interactions. *Trends in Ecology & Evolution* **27:** 705-711.

Franz G, Gencheva E, Kerremans P (1994). Improved stability of genetic sex-separation strains for the Mediterranean fruit fly, *Ceratitis capitata*. *Genome* **37:** 72-82.

Franz G (2005). Genetic sexing strains in Mediterranean fruit fly, an example for other species amenable to large-scale rearing for the sterile insect technique. In: Dyck VA, Hendrichs J, Robinson AS (eds). *Sterile insect technique: Principles and practice in area-wide integrated pest management*. Springer: Dordrecht, The Netherlands. pp 427-451.

Frohman MA, Dush MK, Martin GR (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* **85:** 8998-9002.

Frydman HM, Li JM, Robson DN, Wieschaus E (2006). Somatic stem cell niche tropism in *Wolbachia*. *Nature* **441:** 509-512.

Fu G, Condon KC, Epton MJ, Gong P, Jin L, Condon GC *et al* (2007). Female-specific insect lethality engineered using alternative splicing. *Nature Biotechnology* **25:** 353-357.

Fukatsu T, Hosokawa T (2002). Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, *Megacopta punctatissima*. *Appl Environ Microbiol* **68:** 389-396.

Gabrieli P, Falaguerra A, Siciliano P, Gomulski LM, Scolari F, Zacharopoulou A *et al* (2010). Sex and the single embryo: early development in the Mediterranean fruit fly, *Ceratitis capitata*. *BMC Dev Biol* **10:** 12.

Gilchrist AS, Ling AE (2006). DNA microsatellite analysis of naturally occurring colour intermediates between *Bactrocera tryoni* (Froggatt) and *Bactrocera neohumeralis* (Hardy)(Diptera: Tephritidae). *Aust J Entomol* **45:** 157-162.

Gilchrist AS, Cameron EC, Sved JA, Meats AW (2012). Genetic Consequences of Domestication and Mass Rearing of Pest Fruit Fly *Bactrocera tryoni* (Diptera: Tephritidae). *J Econ Entomol* **105:** 1051-1056.

Gilchrist AS, Meats AW (2012). Factors affecting the dispersal of large-scale releases of the Queensland fruit fly, *Bactrocera tryoni*. *J App Ent* **136:** 252-262.

Gomulski LM, Dimopoulos G, Xi ZY, Soares MB, Bonaldo MF, Malacrida AR *et al* (2008). Gene discovery in an invasive tephritid model pest species, the Mediterranean fruit fly, *Ceratitis capitata*. *BMC Genomics* **9:** 243.

González AN, Lu H, Erickson JW (2008). A shared enhancer controls a temporal switch between promoters during *Drosophila* primary sex determination. *Proc Natl Acad Sci USA* **105:** 18436-18441.

Goralski TJ, Edström JE, Baker BS (1989). The sex determination locus *transformer-2* of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell* **56:** 1011-1018.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I *et al* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29:** 644-652.

Graham P, Penn JK, Schedl P (2003). Masters change, slaves remain. *Bioessays* **25:** 1-4.

Guernier V, Hochberg ME, Guégan J-F (2004). Ecology drives the worldwide distribution of human diseases. *PLoS Biol* **2:** e141.

Guidolin AS, Cônsoli FL (2013). Molecular Characterization of *Wolbachia* Strains Associated with the Invasive Asian Citrus Psyllid *Diaphorina citri* in Brazil. *Microb Ecol* **65:** 475-486.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J *et al* (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8:** 1494-1512.

Hagen KS (1966). Dependence of the olive fly, *Dacus oleae*, larvae on symbiosis with *Pseudomonas savastanoi* for the utilization of olive. *Nature* **209:** 423-424.

Haine ER, Pickup NJ, Cook JM (2005). Horizontal transmission of *Wolbachia* in a *Drosophila* community. *Ecol Entomol* **30:** 464-472.

Hancock DL, Hamacek EL, Lloyd AC, Elson-Harris MM (2000). *The distribution and host plants of fruit flies (Diptera: Tephritidae) in Australia*. Department of Primary Industries, Queensland: Brisbane, Queensland.

Harrison DA, McCoon PE, Binari R, Gilman M, Perrimon N (1998). *Drosophila unpaired* encodes a secreted protein that activates the JAK signaling pathway. *Genes Dev* **12:** 3252-3263.

Heath BD, Butcher RDJ, Whitfield WGF, Hubbard SF (1999). Horizontal transfer of *Wolbachia* between phylogenetically distant insect species by a naturally occurring mechanism. *Curr Biol* **9:** 313-316.

Hedges LM, Brownlie JC, O'Neill SL, Johnson KN (2008). *Wolbachia* and virus protection in insects. *Science* **322:** 702.

Hedley ML, Maniatis T (1991). Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to *tra-2* protein *in vitro*. *Cell* **65:** 579-586.

Heinrich JC, Scott MJ (2000). A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program. *Proc Natl Acad Sci USA* **97:** 8229-8232.

Heinrichs V, Baker BS (1995). The *Drosophila* SR protein RBP1 contributes to the regulation of *doublesex* alternative splicing by recognizing RBP1 RNA target sequences. *EMBO J* **14:** 3987-4000.

Heinrichs V, Ryner LC, Baker BS (1998). Regulation of sex-specific selection of *fruitless* 5' splice sites by *transformer* and *transformer-2*. *Mol Cell Biol* **18:** 450-458.

Hempel LU, Oliver B (2007). Sex-specific *Doublesex*M expression in subsets of *Drosophila* somatic gonad cells. *BMC Dev Biol* **7:** 113.

Hendrichs J, Franz G, Rendon P (1995). Increased effectiveness and applicability of the sterile insect technique through male-only releases for control of Mediterranean fruit flies during fruiting seasons. *J App Ent* **119:** 371-377.

Hertig M, Wolbach SB (1924). Studies on Rickettsia-Like Micro-Organisms in Insects. *J Med Res* **44:** 329-374.

Hilfiker-Kleiner D, Dübendorfer A, Hilfiker A, Nöthiger R (1993). Developmental analysis of two sex-determining genes, M and F, in the housefly, *Musca domestica*. *Genetics* **134:** 1187-1194.

Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH (2008). How many species are infected with *Wolbachia*?--A statistical analysis of current data. *FEMS Microbiol Lett* **281:** 215-220.

Hoffmann AA, Turelli M, Harshman LG (1990). Factors affecting the distribution of cytoplasmic incompatibility in *Drosophila simulans*. *Genetics* **126:** 933-948.

Hoffmann AA, Clancy DJ, Merton E (1994). Cytoplasmic incompatibility in Australian populations of *Drosophila melanogaster*. *Genetics* **136:** 993-999.

Hoffmann AA, Hercus M, Dagher H (1998). Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. *Genetics* **148:** 221-232.

Hoffmann AA, Montgomery BL, Popovici J, Iturbe-Ormaetxe I, Johnson PH, Muzzi F *et al* (2011). Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature* **476:** 454-457.

Hongoh Y, Sharma VK, Prakash T, Noda S, Toh H, Taylor TD *et al* (2008). Genome of an endosymbiont coupling N2 fixation to cellulolysis within protist cells in termite gut. *Science* **322:** 1108-1109.

Horabin JI, Schedl P (1993). Regulated splicing of the *Drosophila sex-lethal* male exon involves a blockage mechanism. *Mol Cell Biol* **13:** 1408-1414.

Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T (2010). *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc Natl Acad Sci USA* **107:** 769-774.

Howard DJ, Bush GL, Breznak JA (1985). The evolutionary significance of bacteria associated with *Rhagoletis*. *Evolution* **39:** 405-417.

Hoy MA, Jeyaprakash A (2005). Microbial diversity in the predatory mite *Metaseiulus occidentalis* (Acari: Phytoseiidae) and its prey, *Tetranychus urticae* (Acari: Tetranychidae). *Biol Control* **32:** 427-441.

Huang Y, Steitz J (2005). SRprises along a messenger's journey. *Molecular cell* **17:** 613-615.

Huigens ME, Luck RF, Klaassen RH, Maas MF, Timmermans MJ, Stouthamer R (2000). Infectious parthenogenesis. *Nature* **405:** 178-179.

Huigens ME, de Almeida RP, Boons PAH, Luck RF, Stouthamer R (2004). Natural interspecific and intraspecific horizontal transfer of parthenogenesis–inducing *Wolbachia* in *Trichogramma* wasps. *Proc R Soc Lond B Biol Sci* **271:** 509-515.

Hurst GDD, Johnson AP, von der Schulenburg JHG, Fuyama Y (2000). Male-killing *Wolbachia* in *Drosophila*: A temperature-sensitive trait with a threshold bacterial density. *Genetics* **156:** 699-709.

Hurst GDD, Jiggins FM (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc R Soc Lond B Biol Sci* **272:** 1525-1534.

Husseneder C, Collier RE (2009). Paratransgenesis in termites. In: Bourtzis K, Miller TA (eds). *Insect Symbiosis*: Boca Raton, FL. pp 361-376.

Inoue K, Hoshijima K, Higuchi I, Sakamoto H, Shimura Y (1992). Binding of the *Drosophila transformer* and *transformer-2* proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. *Proc Natl Acad Sci USA* **89:** 8092-8096.

Jaenike J (2007). Spontaneous emergence of a new *Wolbachia* phenotype. *Evolution* **61:** 2244-2252.

Jamnongluk W, Kittayapong P, Baimai V, O'Neill SL (2002). *Wolbachia* infections of tephritid fruit flies: molecular evidence for five distinct strains in a single host species. *Curr Microbiol* **45:** 255-260.

Jeyaprakash A, Hoy MA (2000). Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. *Insect Mol Biol* **9:** 393-405.

Jolley KA, Chan M-S, Maiden MC (2004). mlstdbNet - distributed multi-locus sequence typing (MLST) databases. *BMC Bioinformatics* **5:** 86.

Jurkevitch E (2011). Riding the Trojan horse: combating pest insects with their own symbionts. *Microbial biotechnology* **4:** 620-627.

Jursnich VA, Burtis KC (1993). A Positive Role in Differentiation for the Male *doublesex* Protein of *Drosophila*. *Dev Biol* **155:** 235-249.

Kageyama D, Traut W (2004). Opposite sex-specific effects of Wolbachia and interference with the sex determination of its host Ostrinia scapulalis. *Proceedings of the Royal Society B: Biological Sciences* **271:** 251-258.

Kaiser W, Huguet E, Casas J, Commin C, Giron D (2010). Plant green-island phenotype induced by leaf-miners is mediated by bacterial symbionts. *Proc R Soc Lond B Biol Sci* **277:** 2311-2319.

Kautz S, Rubin BER, Russell JA, Moreau CS (2013). Surveying the microbiome of ants: Comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. *Appl Environ Microbiol* **79:** 525-534.

Keller GP, Windsor DM, Saucedo JM, Werren JH (2004). Reproductive effects and geographical distributions of two *Wolbachia* strains infecting the Neotropical beetle, *Chelymorpha alternans* Boh. (Chrysomelidae, Cassidinae). *Mol Ecol* **13:** 2405-2420.

Kikuchi Y, Hayatsu M, Hosokawa T, Nagayama A, Tago K, Fukatsu T (2012). Symbiont-mediated insecticide resistance. *Proc Natl Acad Sci USA* **109:** 8618-8622.

Kinnear MW, Bariana HS, Sved JA, Frommer M (1998). Polymorphic microsatellite markers for population analysis of a tephritid pest species, *Bactrocera tryoni*. *Mol Ecol* **7:** 1489-1495.

Kittayapong P, Jamnongluk W, Thipaksorn A, Milne JR, Sindhusake C (2003). *Wolbachia* infection complexity among insects in the tropical rice-field community. *Mol Ecol* **12:** 1049-1060.

Klassen W, Curtis C (2005). History of the sterile insect technique. In: Dyck VA, Hendrichs J, Robinson AS (eds). *Sterile insect technique: Principles and practice in area-wide integrated pest management*. Springer. pp 3-36.

Klasson L, Westberg J, Sapountzis P, Näslund K, Lutnaes Y, Darby AC *et al* (2009). The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*. *Proc Natl Acad Sci USA* **106:** 5725-5730.

Koch H, Schmid-Hempel P (2011). Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc Natl Acad Sci USA* **108:** 19288-19292.

Kounatidis I, Crotti E, Sapountzis P, Sacchi L, Rizzi A, Chouaia B *et al* (2009). *Acetobacter tropicalis* is a major symbiont of the olive fruit fly (*Bactrocera oleae*). *Appl Environ Microbiol* **75:** 3281-3288.

Kraaijeveld K, Reumer BM, Mouton L, Kremer N, Vavre F, Alphen JJM (2011). Does a parthenogenesis-inducing *Wolbachia* induce vestigial cytoplasmic incompatibility? *Naturwissenschaften* **98:** 175-180.

Kramer SG, Jinks TM, Schedl P, Gergen JP (1999). Direct activation of *Sex-lethal* transcription by the *Drosophila runt* protein. *Development* **126:** 191-200.

Kriesner P, Hoffmann AA, Lee SF, Turelli M, Weeks AR (2013). Rapid Sequential Spread of Two *Wolbachia* Variants in *Drosophila simulans*. *PLoS Pathog* **9:** e1003607.

Krosch MN, Schutze MK, Armstrong KF, Graham GC, Yeates DK, Clarke AR (2012). A molecular phylogeny for the Tribe Dacini (Diptera: Tephritidae): Systematic and biogeographic implications. *Mol Phylogenet Evol* **64:** 513-523.

Lagos D, Ruiz MF, Sanchez L, Komitopoulou K (2005). Isolation and characterization of the *Bactrocera oleae* genes orthologous to the sex determining *Sex-lethal* and *doublesex* genes of *Drosophila melanogaster*. *Gene* **348:** 111-121.

Lagos D, Koukidou M, Savakis C, Komitopoulou K (2007). The *transformer* gene in *Bactrocera oleae*: the genetic switch that determines its sex fate. *Insect Mol Biol* **16:** 221-230.

Landmann F, Orsi GA, Loppin B, Sullivan W (2009). *Wolbachia*-mediated cytoplasmic incompatibility is associated with impaired histone deposition in the male pronucleus. *PLoS Pathog* **5:** 215-220.

Langford EA, Nielsen UN, Johnson SN, Riegler M (2014). Susceptibility of Queensland fruit fly, *Bactrocera tryoni* (Froggatt)(Diptera: Tephritidae), to entomopathogenic nematodes. *Biol Control* **69:** 34-39.

Lauzon C, Sjogren R, Prokopy R (2000). Enzymatic capabilities of bacteria associated with apple maggot flies: a postulated role in attraction. *J Chem Ecol* **26:** 953-967.

Le Clec'h W, Braquart-Varnier C, Raimond M, Ferdy JB, Bouchon D, Sicard M (2012). High Virulence of *Wolbachia* after Host Switching: When Autophagy Hurts. *PLoS Pathog* **8:** e1002844.

Lecuit T, Samanta R, Wieschaus E (2002). *slam* Encodes a Developmental Regulator of Polarized Membrane Growth during Cleavage of the *Drosophila* Embryo. *Developmental Cell* **2:** 425-436.

Leroy PD, Sabri A, Heuskin S, Thonart P, Lognay G, Verheggen FJ *et al* (2011). Microorganisms from aphid honeydew attract and enhance the efficacy of natural enemies. *Nature communications* **2:** 348.

Liang H-L, Nien C-Y, Liu H-Y, Metzstein MM, Kirov N, Rushlow C (2008). The zinc-finger protein *Zelda* is a key activator of the early zygotic genome in *Drosophila*. *Nature* **456:** 400-403.

Lindquist D, Abusowa M, Hall M (1992). The New World screwworm fly in Libya: a review of its introduction and eradication. *Med Vet Entomol* **6:** 2-8.

Livak KJ, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25:** 402-408.

Lo N, Paraskevopoulos C, Bourtzis K, O'Neill SL, Werren JH, Bordenstein SR *et al* (2007). Taxonomic status of the intracellular bacterium *Wolbachia pipientis*. *Int J Syst Evol Microbiol* **57:** 654-657.

Lozupone C, Knight R (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71:** 8228-8235.

Maes S, Machtelinckx T, Moens M, Grégoire JC, De Clercq P (2012). The influence of acclimation, endosymbionts and diet on the supercooling capacity of the predatory bug *Macrolophus pygmaeus*. *BioControl* **57:** 643-651.

Marchini D, Rosetto M, Dallai R, Marri L (2002). Bacteria associated with the oesophageal bulb of the medfly *Ceratitis capitata* (Diptera: Tephritidae). *Curr Microbiol* **44:** 120-124.

Martínez H, Toledo J, Liedo P, Mateos M (2012). Survey of Heritable Endosymbionts in Southern Mexico Populations of the Fruit Fly Species *Anastrepha striata* and *A. ludens*. *Curr Microbiol* **65:** 711-718.

Mattick JS, Makunin IV (2006). Non-coding RNA. *Human molecular genetics* **15:** R17-R29.

Mattox W, Palmer MJ, Baker BS (1990). Alternative splicing of the sex determination gene *transformer-2* is sex-specific in the germ line but not in the soma. *Genes Dev* **4:** 789-805.

May AWS (1961). The fruit fly problem in eastern Australia. *Aust J Entomol* **1:** 1-4.

Mazzon L, Piscedda A, Simonato M, Martinez-Sañudo I, Squartini A, Girolami V (2008). Presence of specific symbiotic bacteria in flies of the subfamily Tephritinae (Diptera Tephritidae) and their phylogenetic relationships: proposal of '*Candidatus* Stammerula tephritidis'. *Int J Syst Evol Microbiol* **58:** 1277-1287.

McKeown M, Belote JM, Baker BS (1987). A molecular analysis of *transformer*, a gene in *Drosophila melanogaster* that controls female sexual differentiation. *Cell* **48:** 489-499.

McMeniman CJ, Lane AM, Fong AW, Voronin DA, Iturbe-Ormaetxe I, Yamada R *et al* (2008). Host adaptation of a *Wolbachia* strain after long-term serial passage in mosquito cell lines. *Appl Environ Microbiol* **74:** 6963-6969.

McMeniman CJ, Lane RV, Cass BN, Fong AWC, Sidhu M, Wang YF *et al* (2009). Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti*. *Science* **323:** 141-144.

McMeniman CJ, O'Neill SL (2010). A virulent *Wolbachia* infection decreases the viability of the dengue vector *Aedes aegypti* during periods of embryonic quiescence. *PLoS Neglected Tropical Diseases* **4:** e748.

Meats A (1981). The bioclimatic potential of the Queensland fruit fly, *Dacus tryoni*, in Australia. *Proceedings of the Ecological Society of Australia* **11:** 151-161.

Meats A, Maheswaran P, Frommer M, Sved J (2002). Towards a male-only release system for SIT with the Queensland fruit fly, *Bactrocera tryoni,* using a genetic sexing strain with a temperature-sensitive lethal mutation. *Genetica* **116:** 97-106.

Meats A, Pike N, An X, Raphael K, Wang WY (2003). The effects of selection for early (day) and late (dusk) mating lines of hybrids of *Bactrocera tryoni* and *Bactrocera neohumeralis*. *Genetica* **119:** 283-293.

Meats A, Holmes HM, Kelly GL (2004). Laboratory adaptation of *Bactrocera tryoni* (Diptera: Tephritidae) decreases mating age and increases protein consumption and number of eggs produced per milligram of protein. *Bull Entomol Res* **94:** 517-524.

Meats A (2006). Attributes pertinent to overwintering potential do not explain why *Bactrocera neohumeralis* (Hardy)(Diptera: Tephritidae) does not spread further south within the geographical range of *B. tryoni* (Froggatt). *Aust J Entomol* **45:** 20-25.

Meats A, Edgerton JE (2008). Short- and long-range dispersal of the Queensland fruit fly, *Bactrocera tryoni* and its relevance to invasive potential, sterile insect technique and surveillance trapping. *Australian Journal of Experimental Agriculture* **48:** 1237-1245.

Meats A, Streamer K, Gilchrist AS (2009). Bacteria as food had no effect on fecundity during domestication of the fruit fly, *Bactrocera tryoni*. *J App Ent* **133:** 633-639.

Meise M, Hilfiker-Kleiner D, Dubendorfer A, Brunner C, Nothiger R, Bopp D (1998). *Sex-lethal*, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*. *Development* **125:** 1487-1494.

Miller WJ, Riegler M (2006). Evolutionary dynamics of *w*Au-like *Wolbachia* variants in neotropical *Drosophila* spp. *Appl Environ Microbiol* **72:** 826-835.

Miller WJ, Ehrman L, Schneider D (2010). Infectious speciation revisited: impact of symbiont-depletion on female fitness and mating behavior of *Drosophila paulistorum*. *PLoS Pathog* **6:** e1001214.

Min KT, Benzer S (1997). *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent, causing degeneration and early death. *Proc Natl Acad Sci USA* **94:** 10792-10796.

Miyatake T (1998). Genetic changes of life history and behavioral traits during mass-rearing in the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae). *Res Pop Ecol* **40:** 301-310.

Morag N, Klement E, Saroya Y, Lensky I, Gottlieb Y (2012). Prevalence of the symbiont *Cardinium* in *Culicoides* (Diptera: Ceratopogonidae) vector species is associated with land surface temperature. *FASEB J* **26:** 4025-4034.

Moreira LA, Iturbe-Ormaetxe I, Jeffery JA, Lu G, Pyke AT, Hedges LM *et al* (2009). A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and *Plasmodium*. *Cell* **139:** 1268-1278.

Moreno D, Sanchez M, Robacker D, Worley J (1991). Mating competitiveness of irradiated Mexican fruit fly (Diptera: Tephritidae). *J Econ Entomol* **84:** 1227-1234.

Morrow J, Scott L, Congdon B, Yeates D, Frommer M, Sved J (2000). Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. *Evolution* **54:** 899-910.

Mouton L, Henri H, Bouletreau M, Vavre F (2003). Strain-specific regulation of intracellular *Wolbachia* density in multiply infected insects. *Mol Ecol* **12:** 3459-3465.

Mueller UG, Mikheyev AS, Hong E, Sen R, Warren DL, Solomon SE *et al* (2011). Evolution of cold-tolerant fungal symbionts permits winter fungiculture by leafcutter ants at the northern frontier of a tropical ant-fungus symbiosis. *Proc Natl Acad Sci USA* **108:** 4053-4056.

Murphy KM, MacRae IC, Teakle DS (1988). Nitrogenase activity in the Queensland fruit fly, *Dacus tryoni*. *Aust J Biol Sci* **41:** 447-451.

Murphy KM, Teakle DS, MacRae IC (1994). Kinetics of colonization of adult Queensland fruit flies (*Bactrocera tryoni*) by dinitrogen-fixing alimentary tract bacteria. *Appl Environ Microbiol* **60:** 2508-2517.

Nagoshi RN, McKeown M, Burtis KC, Belote JM, Baker BS (1988). The control of alternative splicing at genes regulating sexual differentiation in *D. melanogaster*. *Cell* **53:** 229-236.

Narita S, Kageyama D, Nomura M, Fukatsu T (2007). Unexpected mechanism of symbiont-induced reversal of insect sex: Feminizing *Wolbachia* continuously acts on the butterfly *Eurema hecabe* during larval development. *Appl Environ Microbiol* **73:** 4332-4341.

Nawrocki EP, Kolbe DL, Eddy SR (2009). Infernal 1.0: inference of RNA alignments. *Bioinformatics* **25:** 1335-1337.

Negri I, Pellecchia M, Mazzoglio PJ, Patetta A, Alma A (2006). Feminizing *Wolbachia* in *Zyginidia pullula* (Insecta, Hemiptera), a leafhopper with an XX/XO sex-determination system. *Proc R Soc Lond B Biol Sci* **273:** 2409-2416.

Negri I, Franchini A, Gonella E, Daffonchio D, Mazzoglio PJ, Mandrioli M *et al* (2009). Unravelling the *Wolbachia* evolutionary role: the reprogramming of the host genomic imprinting. *Proc R Soc Lond B Biol Sci* **276:** 2485-2491.

O'Neill SL, Giordano R, Colbert AM, Karr TL, Robertson HM (1992). 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc Natl Acad Sci USA* **89:** 2699-2702.

Oliver KM, Russell JA, Moran NA, Hunter MS (2003). Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc Natl Acad Sci USA* **100:** 1803-1807.

Osborne R, Meats A, Frommer M, Sved JA, Drew RAI, Robson MK (1997). Australian Distribution of 17 Species of Fruit Flies (Diptera: Tephritidae) Caught in Cue Lure Traps in February 1994. *Aust J Entomol* **36:** 45-50.

Oshima Y, Gotoh Y (1987). Signals for the selection of a splice site in pre-mRNA. Computer analysis of splice junction sequences and like sequences. *J Mol Biol* **195:** 247-259.

Pane A, Salvemini M, Delli Bovi P, Polito C, Saccone G (2002). The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development* **129:** 3715-3725.

Pane A, De Simone A, Saccone G, Polito C (2005). Evolutionary conservation of *Ceratitis capitata transformer* gene function. *Genetics* **171:** 615-624.

Papafotiou G, Oehler S, Savakis C, Bourtzis K (2011). Regulation of *Wolbachia* ankyrin domain encoding genes in *Drosophila* gonads. *Research in Microbiology* **162:** 764-772.

Paraskevopoulos C, Bordenstein SR, Wernegreen JJ, Werren JH, Bourtzis K (2006). Toward a *Wolbachia* multilocus sequence typing system: discrimination of *Wolbachia* strains present in *Drosophila* species. *Curr Microbiol* **53:** 388-395.

Parker A, Mehta K (2007). Sterile insect technique: A model for dose optimization for improved sterile insect quality. *Florida Entomologist* **90:** 88-95.

Paroush Z, Finley RL, Jr., Kidd T, Wainwright SM, Ingham PW, Brent R *et al* (1994). *Groucho* is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79:** 805-815.

Penalva LO, Sanchez L (2003). RNA binding protein *sex-lethal* (*Sxl*) and control of *Drosophila* sex determination and dosage compensation. *Microbiol Mol Biol Rev* **67:** 343-359.

Perrot-Minnot MJ, Guo LR, Werren JH (1996). Single and double infections with *Wolbachia* in the parasitic wasp *Nasonia vitripennis*: effects on compatibility. *Genetics* **143:** 961-972.

Petri L (1910). Untersuchungen uber die Darmbakterien der Olivenfliege. *Zentr Bakt Parasitenk Infekt II* **26:** 357-367.

PHA (2008). Draft National Fruit Fly Strategy. Plant Health Australia: Canberra, Australia.

PHA (2011). The Australian handbook for the identification of fruit flies. In: Australia PH (ed), Version 1.0 edn. Plant Health Australia: Deakin, A.C.T.

Pinto SB, Stainton K, Harris S, Kambris Z, Sutton ER, Bonsall MB *et al* (2013). Transcriptional Regulation of *Culex pipiens* Mosquitoes by *Wolbachia* Influences Cytoplasmic Incompatibility. *PLoS Pathog* **9:** e1003647.

Poinsot D, Bourtzis K, Markakis G, Savakis C, Merçot H (1998). *Wolbachia* transfer from *Drosophila melanogaster* into *D. simulans*: host effect and cytoplasmic incompatibility relationships. *Genetics* **150:** 227-237.

Prabhakar CS, Sood P, Kanwar SS, Sharma PN, Kumar A, Mehta PK (2013). Isolation and characterization of gut bacteria of fruit fly, *Bactrocera tau* (Walker). *Phytoparasitica* **41:** 193-201.

Prosser WA, Douglas AE (1991). The aposymbiotic aphid: An analysis of chlortetracycline-treated pea aphid, *Acyrthosiphon pisum*. *J Insect Physiol* **37:** 713-719.

Pultz MA, Baker BS (1995). The dual role of *hermaphrodite* in the *Drosophila* sex determination regulatory hierarchy. *Development* **121:** 99-111.

Quimio GM, Walter GH (2001). Host preference and host suitability in an egg-pupal fruit fly parasitoid, *Fopius arisanus* (Sonan)(Hym., Braconidae). *J App Ent* **125:** 135-140.

R Core Team (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria.

Raphael KA, Whyard S, Shearman DCA, An X, Frommer M (2004). *Bactrocera tryoni* and closely related pest tephritids--molecular analysis and prospects for transgenic control strategies. *Insect Biochem Mol Biol* **34:** 167-176.

Raphael KA, Shearman DCA, Streamer K, Morrow JL, Handler AM, Frommer M (2011). Germ-line transformation of the Queensland fruit fly, *Bactrocera tryoni*, using a *piggyBac* vector in the presence of endogenous *piggyBac* elements. *Genetica* **139:** 91-97.

Raychoudhury R, Baldo L, Oliveira DC, Werren JH (2009). Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. *Evolution* **63:** 165-183.

Reuter M, Lehmann L, Guillaume F (2008). The spread of incompatibility-inducing parasites in sub-divided host populations. *BMC Evol Biol* **8:** 134.

Reynolds KT, Thomson LJ, Hoffmann AA (2003). The effects of host age, host nuclear background and temperature on phenotypic effects of the virulent *Wolbachia* strain popcorn in *Drosophila melanogaster*. *Genetics* **164:** 1027-1034.

Riegler M, Stauffer C (2002). *Wolbachia* infections and superinfections in cytoplasmically incompatible populations of the European cherry fruit fly *Rhagoletis cerasi* (Diptera, Tephritidae). *Mol Ecol* **11:** 2425-2434.

Riegler M, Charlat S, Stauffer C, Mercot H (2004). *Wolbachia* transfer from *Rhagoletis cerasi* to *Drosophila simulans*: investigating the outcomes of host-symbiont coevolution. *Appl Environ Microbiol* **70:** 273-279.

Riegler M, O'Neill SL (2007). Evolutionary dynamics of insect symbiont associations. *Trends in Ecology & Evolution* **22:** 625-627.

Riegler M, Iturbe-Ormaetxe I, Woolfit M, Miller WJ, O'Neill SL (2012). Tandem repeat markers as novel diagnostic tools for high resolution fingerprinting of *Wolbachia*. *BMC Microbiol* **12:** S12.

Rigaud T, Juchault P (1995). Success and failure of horizontal transfers of feminizing *Wolbachia* endosymbionts in woodlice. *J Evol Biol* **8:** 249-255.

Robinson AS (2002). Genetic sexing strains in medfly, *Ceratitis capitata*, sterile insect technique programmes. *Genetica* **116:** 5-13.

Robinson AS (2005). Genetic basis of the sterile insect technique. In: Dyck VA, Hendrichs J, Robinson AS (eds). *Sterile Insect Technique*. Springer. pp 95-114.

Rocha LS, Mascarenhas RO, Perondini ALP, Selivon D (2005). Occurrence of *Wolbachia* in Brazilian samples of *Ceratitis capitata* (Wiedemann)(Diptera: Tephritidae). *Neotropical Entomology* **34:** 1013-1015.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S *et al* (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61:** 539-542.

Rose LS, Wieschaus E (1992). The *Drosophila* cellularization gene *nullo* produces a blastoderm-specific transcript whose levels respond to the nucleocytoplasmic ratio. *Genes Dev* **6:** 1255-1268.

Rossler Y, Koltin Y (1976). The Genetics of the Mediterranean Fruitfly, *Ceratitis capitata*: Three Morphological Mutations. *Ann Ent Soc Am* **69:** 604-608.

Rothwell WF, Sullivan W (2000). Fluorescent analysis of *Drosophila* embryos. In: Sullivan W, Ashburner M, Hawley RS (eds). *Drosophila Protocols*. Cold Spring Harbor Laboratory Press: Cold Spring Harbour, NY. pp 141-157.

Rousset F, Solignac M (1995). Evolution of single and double *Wolbachia* symbioses during speciation in the *Drosophila simulans* complex. *Proc Natl Acad Sci USA* **92:** 6389-6393.

Royer JE, Hancock DL (2012). New distribution and lure records of Dacinae (Diptera: Tephritidae) from Queensland, Australia, and description of a new species of *Dacus Fabricius*. *Aust J Entomol* **51:** 239-247.

Rozhok A, Bilousov O, Kolodochka L, Zabludovska S, Kozeretska I (2011). Horizontal transmission of intracellular endosymbiotic bacteria: A case between mites and fruit flies and its evolutionary implications. *Dros Inf Serv* **94:** 74-80.

Rugman-Jones PF, Hoddle MS, Stouthamer R (2007). Population genetics of *Scirtothrips perseae*: tracing the origin of a recently introduced exotic pest of Californian avocado orchards, using mitochondrial and microsatellite DNA markers. *Entomol Exp Appl* **124:** 101-115.

Ruiz MF, Stefani RN, Mascarenhas RO, Perondini AL, Selivon D, Sanchez L (2005). The gene *doublesex* of the fruit fly *Anastrepha obliqua* (Diptera, Tephritidae). *Genetics* **171:** 849-854.

Ruiz MF, Milano A, Salvemini M, Eirin-Lopez JM, Perondini ALP, Selivon D *et al* (2007). The Gene *Transformer* of *Anastrepha* Fruit Flies (Diptera, Tephritidae) and Its Evolution in Insects. *PLoS One* **2:** e1239.

Rull J, Encarnacin N, Birke A (2012). Mass rearing history and irradiation affect mating performance of the male fruit fly, *Anastrepha obliqua*. *J Insect Sci* **12:** 45.

Ryner LC, Baker BS (1991). Regulation of *doublesex* pre-mRNA processing occurs by 3'-splice site activation. *Genes Dev* **5:** 2071-2085.

Ryner LC, Goodwin SF, Castrillon DH, Anand A, Villella A, Baker BS *et al* (1996). Control of male sexual behavior and sexual orientation in *Drosophila* by the *fruitless* gene. *Cell* **87:** 1079-1089.

Ryu J-H, Kim S-H, Lee H-Y, Bai JY, Nam Y-D, Bae J-W *et al* (2008). Innate immune homeostasis by the homeobox gene *caudal* and commensal-gut mutualism in *Drosophila*. *Science* **319:** 777-782.

Sacchetti P, Granchietti A, Landini S, Viti C, Giovannetti L, Belcari A (2008). Relationships between the olive fly and bacteria. *J App Ent* **132:** 682-689.

Sacchetti P, Ghiardi B, Granchietti A, Stefanini FM, Belcari A (2014). Development of probiotic diets for the olive fly: evaluation of their effects on fly longevity and fecundity. *Ann Appl Biol* **164:** 138-150.

Saccone G, Peluso I, Artiaco D, Giordano E, Bopp D, Polito LC (1998). The *Ceratitis capitata* homologue of the *Drosophila* sex-determining gene *Sex-lethal* is structurally conserved, but not sex-specifically regulated. *Development* **125:** 1495-1500.

Saccone G, Pane A, Polito LC (2002). Sex determination in flies, fruitflies and butterflies. *Genetica* **116:** 15-23.

Saccone G, Pane A, De Simone A, Salvemini M, Milano A, Annunziata L *et al* (2007). New sexing strains for mediterranean fruit fly *Ceratitis capitata*: *transforming* females into males. *Area-Wide Control of Insect Pests*. Springer. pp 95-102.

Saccone G, Salvemini M, Pane A, Polito LC (2008). Masculinization of XX *Drosophila* transgenic flies expressing the *Ceratitis capitata* DoublesexM isoform. *Int J Dev Biol* **52:** 1051-1057.

Sakamoto H, Inoue K, Higuchi I, Ono Y, Shimura Y (1992). Control of *Drosophila Sex-lethal* pre-mRNA splicing by its own female-specific product. *Nucl Acids Res* **20:** 5533-5540.

Salunke BK, Salunkhe RC, Dhotre DP, Walujkar SA, Khandagale AB, Chaudhari R *et al* (2012). Determination of *Wolbachia* Diversity in Butterflies from Western Ghats, India, by a Multigene Approach. *Appl Environ Microbiol* **78:** 4458-4467.

Salvemini M, Robertson M, Aronson B, Atkinson P, Polito LC, Saccone G (2009). *Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. *Int J Dev Biol* **53:** 109-120.

Salz HK, Erickson JW (2010). Sex determination in *Drosophila*: The view from the top. *Fly* **4:** 60-70.

Sanders JG, Palumbi SR (2011). Populations of *Symbiodinium muscatinei* show strong biogeographic structuring in the intertidal anemone *Anthopleura elegantissima*. *Biol Bull* **220:** 199-208.

Sarakatsanou A, Diamantidis AD, Papanastasiou SA, Bourtzis K, Papadopoulos NT (2011). Effects of *Wolbachia* on fitness of the Mediterranean fruit fly (Diptera: Tephritidae). *J App Ent* **135:** 554-563.

Sarno F, Ruiz MF, Eirín-Lápez JM, Perondini AL, Selivon D, Sànchez L (2010). The gene *transformer-2* of *Anastrepha* fruit flies (Diptera, Tephritidae) and its evolution in insects. *BMC Evol Biol* **10:** 140.

Scarborough CL, Ferrari J, Godfray H (2005). Aphid protected from pathogen by endosymbiont. *Science* **310:** 1781-1781.

Schejter ED, Wieschaus E (1993). *bottleneck* acts as a regulator of the microfilament network governing cellularization of the *Drosophila* embryo. *Cell* **75:** 373-385.

Schetelig MF, Horn C, Handler AM, Wimmer EA (2007). Development of an embryonic lethality system in Mediterranean fruit fly *Ceratitis capitata*. *Area-Wide Control of Insect Pests*. Springer. pp 85-93.

Schetelig MF, Caceres C, Zacharopoulou A, Franz G, Wimmer EA (2009). Conditional embryonic lethality to improve the sterile insect technique in *Ceratitis capitata* (Diptera: Tephritidae). *BMC Biology* **7:** 4.

Schetelig MF, Handler AM (2012a). A transgenic embryonic sexing system for *Anastrepha suspensa* (Diptera: Tephritidae). *Insect Biochem Mol Biol* **42:** 790-795.

Schetelig MF, Handler AM (2012b). Strategy for enhanced transgenic strain development for embryonic conditional lethality in *Anastrepha suspensa*. *Proc Natl Acad Sci USA* **109:** 9348-9353.

Schetelig MF, Milano A, Saccone G, Handler AM (2012). Male only progeny in *Anastrepha suspensa* by RNAi-induced sex reversion of chromosomal females. *Insect Biochem Mol Biol* **42:** 51-57.

Schmittgen TD, Livak KJ (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* **3:** 1101-1108.

Schneider D, Miller WJ, Riegler M (2012). Arthropods Shopping for *Wolbachia*. In: Zchori-Fein E, Bourtzis K (eds). *Manipulative Tenants: Bacteria Associated With Arthropods*. CRC Press: Boca Raton, FL. pp 149-174.

Schuler H, Arthofer W, Krumböck S, Köppler K, Vogt H, Teixeira LAF *et al* (2009). The bacterial endosymbiont *Wolbachia* in the invasive cherry fruit fly *Rhagoletis cingulata* (Diptera, Tephritidae). *Mitt Dt Ges Alig Ang Ento* **17:** 99-101.

Schuler H, Arthofer W, Riegler M, Bertheau C, Krumböck S, Köppler K *et al* (2011). Multiple *Wolbachia* infections in *Rhagoletis pomonella*. *Entomol Exp Appl* **139:** 138-144.

Schuler H, Bertheau C, Egan SP, Feder JL, Riegler M, Schlick-Steiner BC *et al* (2013). Evidence for a recent horizontal transmission and spatial spread of *Wolbachia* from endemic *Rhagoletis cerasi* (Diptera: Tephritidae) to invasive *Rhagoletis cingulata* in Europe. *Mol Ecol* **22:** 4101-4111.

Schüpbach T (1982). Autosomal mutations that interfere with sex determination in somatic cells of *Drosophila* have no direct effect on the germline. *Dev Biol* **89:** 117-127.

Schutt C, Nothiger R (2000). Structure, function and evolution of sex-determining systems in Dipteran insects. *Development* **127:** 667.

Serbus LR, Casper-Lindley C, Landmann F, Sullivan W (2008). The genetics and cell biology of *Wolbachia*-host interactions. *Annu Rev Genet* **42:** 683-707.

Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I, Rosenberg E (2010). Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **107:** 20051-20056.

Shearman DCA, Frommer M (1998). The *Bactrocera tryoni* homologue of the *Drosophila melanogaster* sex-determination gene *doublesex*. *Insect Mol Biol* **7:** 355-366.

Shearman DCA (2002). The evolution of sex determination systems in dipteran insects other than *Drosophila*. *Genetica* **116:** 25-43.

Shearman DCA, Frommer M, Morrow JL, Raphael KA, Gilchrist AS (2010). Interspecific Hybridization as a Source of Novel Genetic Markers for the Sterile Insect Technique in *Bactrocera tryoni* (Diptera: Tephritidae). *J Econ Entomol* **103:** 1071-1079.

Shen GM, Dou W, Huang Y, Jiang XZ, Smagghe G, Wang JJ (2013). *In silico* cloning and annotation of genes involved in the digestion, detoxification and RNA interference mechanism in the midgut of *Bactrocera dorsalis* [Hendel (Diptera: Tephritidae)]. *Insect Mol Biol* **22:** 354-365.

Sievert V, Kuhn S, Paululat A, Traut W (2000). Sequence conservation and expression of the *Sex-lethal* homologue in the fly *Megaselia scalaris*. *Genome* **43:** 382-390.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Ent Soc Am* **87:** 651-701.

Sintupachee S, Milne JR, Poonchaisri S, Baimai V, Kittayapong P (2006). Closely Related *Wolbachia* Strains within the Pumpkin Arthropod Community and the Potential for Horizontal Transmission via the Plant. *Microb Ecol* **51:** 294-301.

Smith PH (1979). Genetic manipulation of the circadian clock's timing of sexual behaviour in the Queensland fruit flies, *Dacus tryoni* and *Dacus neohumeralis*. *Physiol Entomol* **4:** 71-78.

Snowball GJ, Wilson F, Campbell TG, Lukins RG (1962). The utilization of parasites of oriental fruit fly (*Dacus dorsalis*) against Queensland fruit fly (*Strumeta tryoni*). *Crop & Pasture Science* **13:** 443-447.

Sosnowski BA, Belote JM, McKeown M (1989). Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell* **58:** 449-459.

Souza HML, Matioli SR, Souza WN (1988). The adaptation process of *Ceratitis capitata* to the laboratory analysis of life-history traits. *Entomol Exp Appl* **49:** 195-201.

Spinner JE, Cowling AM, Gurr GM, Jessup AJ, Reynolds OL (2011). Parasitoid fauna of Queensland fruit fly, *Bactrocera tryoni* Froggatt (Diptera: Tephritidae) in inland New South Wales, Australia and their potential for use in augmentative biological control. *Aust J Entomol* **50:** 445-452.

Stahlhut JK, Desjardins CA, Clark ME, Baldo L, Russell JA, Werren JH *et al* (2010). The mushroom habitat as an ecological arena for global exchange of *Wolbachia*. *Mol Ecol* **19:** 1940-1952.

Starr DJ, Cline TW (2002). A host parasite interaction rescues *Drosophila* oogenesis defects. *Nature* **418:** 76-79.

Staubach F, Baines JF, Kuenzel S, Bik EM, Petrov DA (2013). Bacterial diversity associated with *Drosophila* in the laboratory and in the natural environment. *PLoS One* **8:** e70749.

Stein JA, Broihier HT, Moore LA, Lehmann R (2002). *Slow as Molasses* is required for polarized membrane growth and germ cell migration in *Drosophila*. *Development* **129:** 3925-3934.

Stouthamer R, Luck RF, Hamilton WD (1990). Antibiotics cause parthenogenetic *Trichogramma* (Hymenoptera/Trichogrammatidae) to revert to sex. *Proc Natl Acad Sci USA* **87:** 2424-2427.

Sun S, Cline TW (2009). Effects of *Wolbachia* infection and ovarian tumor mutations on *Sex-lethal* germline functioning in *Drosophila*. *Genetics* **181:** 1291-1301.

Sun X, Cui L, Li Z (2007). Diversity and phylogeny of *Wolbachia* infecting *Bactrocera dorsalis* (Diptera: Tephritidae) populations from China. *Environ Entomol* **36:** 1283-1289.

Sutherst RW, Collyer BS, Yonow T (2000). The vulnerability of Australian horticulture to the Queensland fruit fly, *Bactrocera* (*Dacus*) *tryoni*, under climate change. *Aust J Agric Res* **51:** 467-480.

Tadros W, Goldman AL, Babak T, Menzies F, Vardy L, Orr-Weaver T *et al* (2007). SMAUG Is a Major Regulator of Maternal mRNA Destabilization in *Drosophila* and Its Translation Is Activated by the PAN GU Kinase. *Developmental Cell* **12:** 143-155.

Tadros W, Lipshitz HD (2009). The maternal-to-zygotic transition: a play in two acts. *Development* **136:** 3033-3042.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28:** 2731-2739.

Tay ST (2013). *Wolbachia* endosymbionts, *Rickettsia felis* and *Bartonella* species, in *Ctenocephalides felis* fleas in a tropical region. *J Vect Ecol* **38:** 200-202.

Teixeira L, Ferreira Á, Ashburner M (2008). The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol* **6:** e1000002.

ten Bosch JR, Benavides JA, Cline TW (2006). The TAGteam DNA motif controls the timing of *Drosophila* pre-blastoderm transcription. *Development* **133:** 1967-1977.

Thaochan N, Drew RAI, Hughes JM, Vijaysegaran S, Chinajariyawong A (2010). Alimentary tract bacteria isolated and identified with API-20E and molecular cloning techniques from Australian tropical fruit flies, *Bactrocera cacuminata* and *B. tryoni*. *J Insect Sci* **10:** 131.

Thaochan N, Chinajariyawong A (2011). Attraction of *Bactrocera cucurbitae* and *B. papayae* (Diptera: Tephritidae) to the Odor of the Bacterium *Enterobacter cloacae*. *The Philippine Agricultural Scientist* **94:** 83-87.

Thomas DD, Donnelly CA, Wood RJ, Alphey LS (2000). Insect population control using a dominant, repressible, lethal genetic system. *Science* **287:** 2474-2476.

Tian H, Peng H, Yao Q, Chen H, Xie Q, Tang B *et al* (2009). Developmental control of a lepidopteran pest *Spodoptera exigua* by ingestion of bacteria expressing dsRNA of a non-midgut gene. *PLoS One* **4:** e6225.

Tian M, Maniatis T (1993). A splicing enhancer complex controls alternative splicing of *doublesex* pre-mRNA. *Cell* **74:** 105-114.

Toledo J, Rull J, Oropeza A, Hernández E, Liedo P (2004). Irradiation of *Anastrepha obliqua* (Diptera: Tephritidae) revisited: optimizing sterility induction. *J Econ Entomol* **97:** 383-389.

Torres M, Sanchez L (1989). The *scute* (T4) gene acts as a numerator element of the X:a signal that determines the state of activity of *sex-lethal* in *Drosophila*. *EMBO J* **8:** 3079-3086.

Torres M, Sanchez L (1992). The segmentation gene *runt* is needed to activate *Sex-lethal,* a gene that controls sex determination and dosage compensation in *Drosophila*. *Genetics Research* **59:** 189-198.

Tram U, Sullivan W (2002). Role of delayed nuclear envelope breakdown and mitosis in *Wolbachia*-induced cytoplasmic incompatibility. *Science* **296:** 1124.

Tsiropoulos GJ (1981). Effect of antibiotics incorporated into defined adult diets on survival and reproduction of the walnut husk fly, *Rhagoletis completa* Cress.(Dipt., Trypetidae). *Zeitschrift für Angewandte Entomologie* **91:** 100-106.

Tsuchida T, Koga R, Fukatsu T (2009). Endosymbiont that broadens food plant range of host insect. In: Bourtzis K, Miller TA (eds). *Insect Symbiosis*. pp 123-142.

Turelli M, Hoffmann AA (1991). Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353:** 440-442.

Tychsen PH, Fletcher BS (1971). Studies on the rhythm of mating in the Queensland fruit fly, *Dacus tryoni*. *J Insect Physiol* **17:** 2139-2156.

Unckless RL, Boelio LM, Herren JK, Jaenike J (2009). *Wolbachia* as populations within individual insects: causes and consequences of density variation in natural populations. *Proc R Soc Lond B Biol Sci* **276:** 2805-2811.

Valcarcel J, Singh R, Zamore PD, Green MR (1993). The protein *Sex-lethal* antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. *Nature* **362:** 171-175.

Vandekerckhove TTM, Watteyne S, Bonne W, Vanacker D, Devaere S, Rumes B *et al* (2003). Evolutionary trends in feminization and intersexuality in woodlice

(Crustacea, Isopoda) infected with *Wolbachia pipientis* (α-Proteobacteria). *Belgian Journal of Zoology* **133:** 61-69.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A *et al* (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3:** research0034-research0034.0011.

Vavre F, Fleury F, Lepetit D, Fouillet P, Boulétreau M (1999). Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. *Mol Biol Evol* **16:** 1711-1723.

Veneti Z, Bentley JK, Koana T, Braig HR, Hurst GDD (2005). A functional dosage compensation complex required for male killing in *Drosophila*. *Science* **307:** 1461-1463.

Walker T, Klasson L, Sebaihia M, Sanders MJ, Thomson NR, Parkhill J *et al* (2007). Ankyrin repeat domain-encoding genes in the *w*Pip strain of *Wolbachia* from the *Culex pipiens* group. *BMC Biology* **5:** 39.

Walker T, Johnson PH, Moreira LA, Iturbe-Ormaetxe I, Frentiu FD, McMeniman CJ *et al* (2011). The *w*Mel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature* **476:** 450-453.

Wang H, Jin L, Zhang H (2011). Comparison of the diversity of the bacterial communities in the intestinal tract of adult *Bactrocera dorsalis* from three different populations. *J Applied Microbiol* **110:** 1390-1401.

Wang H, Jin L, Peng T, Zhang H, Chen Q, Hua Y (2013). Identification of cultivable bacteria in the intestinal tract of *Bactrocera dorsalis* from three different populations and determination of their attractive potential. *Pest Manag Sci* **70:** 80-87.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73:** 5261-5267.

Wang Y, Yu H, Raphael K, Gilchrist AS (2003). Genetic delineation of sibling species of the pest fruit fly *Bactrocera* (Diptera: Tephritidae) using microsatellites. *Bull Entomol Res* **93:** 351-360.

Wang Y, Gilbreath TM, Kukutla P, Yan G, Xu J (2011). Dynamic gut microbiome across life history of the malaria mosquito Anopheles gambiae in Kenya. *PLoS One* **6:** e24767.

Wang Y, Naumann U, Wright ST, Warton DI (2012). mvabund– an R package for model-based analysis of multivariate abundance data. *Methods in Ecology and Evolution* **3:** 471-474.

Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT *et al* (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450:** 560-565.

Warren A (2013). Putting a stop to the rot. In: Morgan H (ed). CSIRO.

Watanabe M, Miura K, Hunter MS, Wajnberg E (2011). Superinfection of cytoplasmic incompatibility-inducing *Wolbachia* is not additive in *Orius strigicollis* (Hemiptera: Anthocoridae). *Heredity* **106:** 642-648.

Weeks AR, Breeuwer JAJ (2001). *Wolbachia*-induced parthenogenesis in a genus of phytophagous mites. *Proc R Soc Lond B Biol Sci* **268:** 2245-2251.

Weeks AR, Turelli M, Harcombe WR, Reynolds KT, Hoffmann AA (2007). From parasite to mutualist: rapid evolution of *Wolbachia* in natural populations of *Drosophila*. *PLoS Biol* **5:** e114.

Weiss B, Aksoy S (2011). Microbiome influences on insect host vector competence. *Trends Parasitol* **27:** 514-522.

Weldon CW (2005). Mass-rearing and sterilisation alter mating behaviour of male Queensland fruit fly, *Bactrocera tryoni* (Froggatt)(Diptera: Tephritidae). *Aust J Entomol* **44:** 158-163.

Werren JH, Windsor DM, Guo LR (1995a). Distribution of *Wolbachia* among neotropical arthropods. *Proc R Soc Lond B Biol Sci* **262:** 197-204.

Werren JH, Zhang W, Guo LR (1995b). Evolution and phylogeny of *Wolbachia*: reproductive parasites of arthropods. *Proc R Soc Lond B Biol Sci* **261:** 55-63.

Werren JH (1997a). *Wolbachia* run amok. *Proc Natl Acad Sci USA* **94:** 11154-11155.

Werren JH (1997b). Biology of *Wolbachia*. *Annu Rev Entomol* **42:** 587-609.

Werren JH, Windsor DM (2000). *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? *Proc R Soc Lond B Biol Sci* **267:** 1277-1285.

Werren JH, Baldo L, Clark ME (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol* **6:** 741-751.

West SA, Cook JM, Werren JH, Godfray HC (1998). *Wolbachia* in two insect host-parasitoid communities. *Mol Ecol* **7:** 1457-1465.

White IM, Elson-Harris MM (1992). *Fruit flies of economic significance: their identification and bionomics*. CAB International: Oxon, UK.

Whyard S, Singh AD, Wong S (2009). Ingested double-stranded RNAs can act as species-specific insecticides. *Insect Biochem Mol Biol* **39:** 824-832.

Willhoeft U, Franz G (1996a). Comparison of the mitotic karyotypes of *Ceratitis capitata*, *Ceratitis rosa*, and *Trirhithrum coffeae* (Diptera: Tephritidae) by C-banding and FISH. *Genome* **39:** 884-889.

Willhoeft U, Franz G (1996b). Identification of the sex-determining region of the *Ceratitis capitata* Y chromosome by deletion mapping. *Genetics* **144:** 737-745.

Wong ACN, Ng P, Douglas AE (2011). Low-diversity bacterial community in the gut of the fruitfly *Drosophila melanogaster*. *Environ Microbiol* **13:** 1889-1900.

Wong ACN, Chaston JM, Douglas AE (2013). The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *ISME J* **7:** 1922-1932.

Xi Z, Khoo CCH, Dobson SL (2006). Interspecific transfer of *Wolbachia* into the mosquito disease vector *Aedes albopictus*. *Proc R Soc Lond B Biol Sci* **273:** 1317-1322.

Yang D, Lu H, Hong Y, Jinks TM, Estes PA, Erickson JW (2001). Interpretation of X chromosome dose at *Sex-lethal* requires non-E-box sites for the basic helix-loop-helix proteins SISB and *daughterless*. *Mol Cell Biol* **21:** 1581-1592.

Yanowitz JL, Deshpande G, Calhoun G, Schedl PD (1999). An N-terminal truncation uncouples the sex-transforming and dosage compensation functions of *Sex-lethal*. *Mol Cell Biol* **19:** 3018.

Yonow T, Sutherst RW (1998). The geographical distribution of the Queensland fruit fly, *Bactrocera (Dacus) tryoni,* in relation to climate. *Aust J Agric Res* **49:** 935-954.

Yu Y, Lee C, Kim J, Hwang S (2005). Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol Bioeng* **89:** 670-679.

Zabalou S, Riegler M, Theodorakopoulou M, Stauffer C, Savakis C, Bourtzis K (2004). *Wolbachia*-induced cytoplasmic incompatibility as a means for insect pest population control. *Proc Natl Acad Sci USA* **101:** 15042-15045.

Zabalou S, Apostolaki A, Pattas S, Veneti Z, Paraskevopoulos C, Livadaras I *et al* (2008). Multiple rescue factors within a *Wolbachia* strain. *Genetics* **178:** 2145-2160.

Zabalou S, Apostolaki A, Livadaras I, Franz G, Robinson AS, Savakis C *et al* (2009). Incompatible insect technique: Incompatible males from a *Ceratitis capitata* genetic sexing strain. *Entomol Exp Appl* **132:** 232-240.

Zhao JT, Frommer M, Sved JA, Zacharopoulou A (1998). Mitotic and polytene chromosome analyses in the Queensland fruit fly, *Bactrocera tryoni* (Diptera: Tephritidae). *Genome* **41:** 510-526.

Zhao JT, Bennett CL, Stewart GJ, Frommer M, Raphael KA (2003a). The *scarlet* eye colour gene of the tephritid fruit fly: *Bactrocera tryoni* and the nature of two eye colour mutations. *Insect Mol Biol* **12:** 263-269.

Zhao JT, Frommer M, Sved JA, Gillies CB (2003b). Genetic and molecular markers of the Queensland fruit fly, *Bactrocera tryoni*. *Journal of Heredity* **94:** 416-420.

Zhou W, Rousset F, O'Neill SL (1998). Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc R Soc Lond B Biol Sci* **265:** 509-515.

Zug R, Hammerstein P (2012). Still a host of hosts for *Wolbachia*: Analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PLoS One* **7:** e38544.

Zug R, Koehncke A, Hammerstein P (2012). Epidemiology in evolutionary time: The case of *Wolbachia* horizontal transmission between arthropod host species. *J Evol Biol* **25:** 2149-2160.

# Appendices

# Appendix A

**Table A.1** Collection data for field-caught fruit flies from this study. Information for collection location, climate and year, specimen specific ID no., gender, result for *wsp* and 16S rDNA screening and GenBank accession number of *COI* partial gene are included.

| | Location | | | | | Climate | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
| *Bactrocera visenda* | Pajinka | 10.42 | 142.32 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 211 | n | n | y |
| *Bactrocera visenda* | Pajinka | 10.42 | 142.32 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 212 | n | n | y |
| *Bactrocera manskii* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A927 | M | 63 | n | n | y |
| *Bactrocera manskii* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A927 | M | 64 | n | n | y |
| *Bactrocera manskii* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A927 | M | 65 | n | n | KC581383 |
| *Bactrocera perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A932 | M | 74 | y | y | y |
| *Bactrocera perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A932 | M | 75 | y | y | y |
| *Bactrocera perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A932 | M | 76 | y | y | y |
| *Bactrocera perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A932 | M | 261 | y | y | y |
| *Bactrocera perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A932 | M | 263 | y | y | y |
| *Bactrocera fallacis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 200 | n | n | KC581382 |
| *Bactrocera fallacis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 201 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 202 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 203 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 204 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 213 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 214 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 215 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 216 | n | n | y |
| *Bactrocera neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 217 | n | n | y |
| *Bactrocera strigifinis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 198 | n | n | y |
| *Bactrocera strigifinis* | Lockerbie | 10.47 | 142.27 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 199 | n | n | y |
| *Bactrocera visenda* | Injinoo | 10.54 | 142.29 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 207 | n | n | KC581404 |
| *Bactrocera visenda* | Injinoo | 10.54 | 142.29 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 208 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera bryoniae* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A921 | M | 56 | n | n | Y |
| *Bactrocera bryoniae* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A921 | M | 57 | n | n | y |
| *Bactrocera jarvisi* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A926 | M | 60 | n | n | y |
| *Bactrocera jarvisi* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A926 | M | 61 | n | n | y |
| *Bactrocera jarvisi* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A926 | M | 62 | n | n | y |
| *Bactrocera murrayi* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A928 | M | 66 | n | n | y |
| *Bactrocera murrayi* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A928 | M | 67 | n | n | y |
| *Bactrocera murrayi* | Torres Strait | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A928 | M | 68 | n | n | y |
| *Bactrocera papayae* | Torres Strait | 10.58 | 142.22 | 1998 | Equatorial | 1843 | 1734 | 30.2 | 29.1 | A930 | M | 70 | n | n | y |
| *Bactrocera peninsularis* | Cape York | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A931 | M | 71 | n | y | KC581392 |
| *Bactrocera peninsularis* | Cape York | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A931 | M | 72 | n | y | y |
| *Bactrocera peninsularis* | Cape York | 10.58 | 142.22 | 1997 | Equatorial | 1843 | 1391 | 30.2 | 29.1 | A931 | M | 73 | n | y | y |
| *Dacus axanus* | Thursday Island | 10.58 | 142.22 | 1998 | Equatorial | 1843 | 1734 | 30.2 | 30.3 | A936 | M | 86 | n | n | y |
| *Dacus axanus* | Thursday Island | 10.58 | 142.22 | 1998 | Equatorial | 1843 | 1734 | 30.2 | 30.3 | A936 | M | 87 | n | n | y |
| *Dacus bellulus* | Thursday Island | 10.58 | 142.22 | 1998 | Equatorial | 1843 | 1734 | 30.2 | 30.3 | A939 | M | 89 | n | n | KC581410 |
| *Dacus bellulus* | Thursday Island | 10.58 | 142.22 | 1998 | Equatorial | 1843 | 1734 | 30.2 | 30.3 | A939 | M | 90 | n | n | y |
| *Dacus bellulus* | Thursday Island | 10.58 | 142.22 | 1998 | Equatorial | 1843 | 1734 | 30.2 | 30.3 | A939 | M | 91 | n | n | y |
| *Bactrocera decurtans* | Seisia | 10.85 | 142.35 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A935 | M | 83 | n | n | KC581376 |
| *Bactrocera decurtans* | Seisia | 10.85 | 142.35 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A935 | M | 84 | n | n | y |
| *Bactrocera decurtans* | Seisia | 10.85 | 142.35 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A935 | M | 85 | y | y | KC581377 |
| *Bactrocera decurtans* | Seisia | 10.85 | 142.35 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A935 | M | 264 | n | n | y |
| *Bactrocera decurtans* | Seisia | 10.85 | 142.35 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A935 | M | 265 | n | n | y |
| *Bactrocera decurtans* | Seisia | 10.85 | 142.35 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A935 | M | 266 | n | n | y |
| *Bactrocera frauenfeldi* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 191 | n | n | KC581380 |
| *Bactrocera frauenfeldi* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 192 | n | n | y |
| *Bactrocera neohumeralis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 193 | n | n | y |
| *Bactrocera neohumeralis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 194 | n | n | y |
| *Bactrocera neohumeralis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 195 | n | n | y |
| *Bactrocera peninsularis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 196 | n | y | KC581393 |
| *Bactrocera peninsularis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 197 | n | y | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera strigifinis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 189 | n | n | y |
| *Bactrocera strigifinis* | Seisia | 10.85 | 142.35 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 190 | n | n | y |
| *Bactrocera fallacis* | New Mapoon | 10.87 | 142.39 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 365 | n | n | y |
| *Bactrocera fallacis* | New Mapoon | 10.87 | 142.39 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 366 | n | n | y |
| *Bactrocera fallacis* | New Mapoon | 10.87 | 142.39 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 367 | n | n | y |
| *Bactrocera visenda* | New Mapoon | 10.87 | 142.39 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 209 | n | n | KC581403 |
| *Bactrocera visenda* | New Mapoon | 10.87 | 142.39 | 2012 | Equatorial | 1843 | 1600 | 30.2 | 30.8 | SC2012 | M | 210 | n | n | y |
| *Bactrocera fallacis* | Bamaga | 10.89 | 142.39 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 205 | n | n | y |
| *Bactrocera fallacis* | Bamaga | 10.89 | 142.39 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 206 | n | n | y |
| *Bactrocera strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A933 | M | 80 | n | n | y |
| *Bactrocera strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A933 | M | 81 | y | y | KC581396 |
| *Bactrocera strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A933 | M | 82 | n | n | y |
| *Bactrocera strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A933 | M | 269 | y | y | y |
| *Bactrocera strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A933 | M | 270 | n | n | y |
| *Bactrocera strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | Equatorial | 1830 | 2031 | 30.2 | 30.3 | A933 | M | 271 | n | n | y |
| *Bactrocera frauenfeldi* | Umagico | 10.92 | 142.39 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 319 | n | n | y |
| *Bactrocera frauenfeldi* | Umagico | 10.92 | 142.39 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 320 | n | n | y |
| *Bactrocera frauenfeldi* | Umagico | 10.92 | 142.39 | 2012 | Equatorial | 1830 | 1600 | 30.2 | 30.8 | SC2012 | M | 321 | n | n | y |
| *Bactrocera allwoodi* | Nhulunbuy | 12.19 | 136.76 | 1997 | Tropical | 1306 | 683 | 30.6 | 30.4 | 97/25.1 | M | 16 | n | n | KC581371 |
| *Bactrocera jarvisi* | Nhulunbuy | 12.19 | 136.76 | 1997 | Tropical | 1306 | 683 | 30.6 | 30.4 | 97/25.1 | M | 15 | n | n | y |
| *Bactrocera bryoniae* | Casuarina | 12.37 | 130.87 | 1998 | Tropical | 1830 | 2509 | 32 | 32.9 | 98/14 | M | 36 | n | n | y |
| *Bactrocera tryoni* | Wanguri | 12.37 | 130.89 | 1998 | Tropical | 1830 | 2509 | 32 | 32.9 | 98/4 | M | 47 | n | n | y |
| *Bactrocera tryoni* | Wanguri | 12.37 | 130.89 | 1998 | Tropical | 1830 | 2509 | 32 | 32.9 | 98/4 | M | 48 | n | n | y |
| *Bactrocera bryoniae* | Howard Springs | 12.47 | 131.06 | 1998 | Tropical | 1897 | 2881 | 32 | 32.9 | 98/13 | M | 37 | n | n | y |
| *Bactrocera bryoniae* | Howard Springs | 12.47 | 131.06 | 1998 | Tropical | 1897 | 2881 | 32 | 32.9 | 98/13 | M | 38 | n | n | KC581373 |
| *Bactrocera tryoni* | Palmerston | 12.48 | 130.98 | 1998 | Tropical | 1910 | 1969 | 32 | 32.9 | 98/10A | M | 116 | n | n | y |
| *Bactrocera tryoni* | Palmerston | 12.48 | 130.98 | 1998 | Tropical | 1910 | 1969 | 32 | 32.9 | 98/10A | M | 117 | n | n | y |
| *Bactrocera tryoni* | Palmerston | 12.48 | 130.98 | 1998 | Tropical | 1910 | 1969 | 32 | 32.9 | 98/10A | M | 118 | n | n | y |
| *Bactrocera tryoni* | Palmerston | 12.48 | 130.98 | 1998 | Tropical | 1910 | 1969 | 32 | 32.9 | 98/10A | M | 119 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 496 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera frauenfeldi* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 497 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 498 | n | n | y |
| *Bactrocera strigifinis* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 514 | n | n | y |
| *Bactrocera strigifinis* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 515 | n | n | y |
| *Bactrocera strigifinis* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 516 | n | n | y |
| *Bactrocera strigifinis* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 517 | n | n | y |
| *Bactrocera strigifinis* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 518 | n | n | y |
| *Bactrocera tryoni* | Lockhart River | 12.60 | 143.41 | 2013 | Equatorial | 2081 | 1325 | 29.8 | 30.1 | JR2013 | M | 561 | n | n | y |
| *Bactrocera bryoniae* | Weipa | 12.63 | 141.92 | 1997 | Equatorial | 1781 | 2496 | 32.7 | 32.6 | A924 | M | 58 | n | n | y |
| *Bactrocera bryoniae* | Weipa | 12.63 | 141.92 | 1997 | Equatorial | 1781 | 2496 | 32.7 | 32.6 | A924 | M | 59 | n | n | y |
| *Bactrocera bryoniae* | Weipa | 12.63 | 141.92 | 1997 | Equatorial | 1781 | 2496 | 32.7 | 32.6 | A924 | M | 272 | n | n | y |
| *Bactrocera bryoniae* | Weipa | 12.63 | 141.92 | 1997 | Equatorial | 1781 | 2496 | 32.7 | 32.6 | A924 | M | 273 | n | n | y |
| *Bactrocera bryoniae* | Weipa | 12.63 | 141.92 | 1997 | Equatorial | 1781 | 2496 | 32.7 | 32.6 | A924 | M | 274 | n | n | y |
| *Bactrocera frauenfeldi* | Weipa | 12.63 | 141.92 | 1998 | Equatorial | 1781 | 1978 | 32.7 | 33 | 98/240a | M | 45 | n | n | y |
| *Bactrocera frauenfeldi* | Weipa | 12.63 | 141.92 | 1998 | Equatorial | 1781 | 1978 | 32.7 | 33 | 98/240a | M | 46 | n | n | y |
| *Bactrocera frauenfeldi* | Weipa | 12.63 | 141.92 | 2001 | Equatorial | 1781 | 1895 | 32.7 | 33 | 01/240A1 | M | 137 | n | n | y |
| *Bactrocera strigifinis* | Weipa | 12.63 | 141.92 | 2001 | Equatorial | 1781 | 1895 | 32.7 | 33 | 01/240A1 | M | 128 | n | n | y |
| *Bactrocera strigifinis* | Weipa | 12.63 | 141.92 | 2001 | Equatorial | 1781 | 1895 | 32.7 | 33 | 01/239B.1 | M | 129 | n | n | y |
| *Bactrocera tryoni* | Weipa | 12.63 | 141.92 | 1998 | Equatorial | 1781 | 1978 | 32.7 | 33 | 98/240a | M | 283 | n | n | y |
| *Dacus axanus* | Weipa | 12.63 | 141.92 | 1998 | Equatorial | 1781 | 1978 | 32.7 | 33 | A937 | M | 88 | y | y | KC581409 |
| *Dacus axanus* | Weipa | 12.63 | 141.92 | 1998 | Equatorial | 1781 | 1978 | 32.7 | 33 | A937 | M | 267 | n | n | y |
| *Dacus axanus* | Weipa | 12.63 | 141.92 | 2001 | Equatorial | 1781 | 1895 | 32.7 | 33 | 01/240A1 | M | 268 | n | n | y |
| *Bactrocera bryoniae* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 535 | n | n | y |
| *Bactrocera bryoniae* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 536 | y | y | KC581372 |
| *Bactrocera bryoniae* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 537 | n | n | y |
| *Bactrocera bryoniae* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 538 | n | n | y |
| *Bactrocera bryoniae* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 539 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 499 | n | n | y |
| *Bactrocera tryoni* | Lockhart River | 12.79 | 143.34 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 560 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.80 | 143.33 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 493 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera frauenfeldi* | Lockhart River | 12.80 | 143.33 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 494 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.80 | 143.33 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 495 | n | n | y |
| *Bactrocera bryoniae* | Lockhart River | 12.80 | 143.32 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 576 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.80 | 143.32 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 500 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.80 | 143.32 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 558 | n | n | y |
| *Bactrocera frauenfeldi* | Lockhart River | 12.80 | 143.32 | 2013 | Equatorial | 2081 | 1327 | 29.8 | 30.24 | JR2013 | M | 559 | n | n | y |
| *Bactrocera bryoniae* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 555 | n | n | y |
| *Bactrocera bryoniae* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 556 | n | n | y |
| *Bactrocera bryoniae* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 557 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 567 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 568 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 569 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 570 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.92 | 143.19 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 571 | n | n | y |
| *Bactrocera strigifinis* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 509 | n | n | y |
| *Bactrocera strigifinis* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 510 | n | n | y |
| *Bactrocera strigifinis* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 511 | n | n | y |
| *Bactrocera strigifinis* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 512 | n | n | y |
| *Bactrocera strigifinis* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 513 | n | n | y |
| *Bactrocera bryoniae* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 554 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 562 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 563 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 564 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 565 | n | n | y |
| *Bactrocera tryoni* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 566 | n | n | y |
| *Dacus axanus* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 572 | n | n | y |
| *Dacus axanus* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 573 | n | n | y |
| *Dacus axanus* | Coen | 13.94 | 143.20 | 2013 | Equatorial | 1174 | 1035 | 31.6 | 31.9 | JR2013 | M | 574 | n | n | y |
| *Dacus axanus* | Cairns | 16.90 | 145.74 | 1996 | Tropical | 2014 | 2083 | 29 | 29.3 | A741 | M | 391 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 479 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 480 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 540 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 541 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 542 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 543 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 544 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 545 | y | y | KC581372 |
| *Bactrocera bryoniae* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 546 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 483 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 484 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 485 | y | y | y |
| *Bactrocera frauenfeldi* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 486 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 487 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.78 | 145.68 | 2012 | Tropical | 2133 | 2203 | 29 | 29.2 | JR2012 | M | 218 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.78 | 145.68 | 2012 | Tropical | 2133 | 2203 | 29 | 29.2 | JR2012 | M | 219 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.78 | 145.68 | 2012 | Tropical | 2133 | 2203 | 29 | 29.2 | JR2012 | M | 242 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.78 | 145.68 | 2012 | Tropical | 2133 | 2203 | 29 | 29.2 | JR2012 | M | 243 | y | y | KC775785 |
| *Bactrocera neohumeralis* | Cairns | 16.78 | 145.68 | 2012 | Tropical | 2133 | 2203 | 29 | 29.2 | JR2012 | M | 244 | y | y | KC775786 |
| *Bactrocera neohumeralis* | Cairns | 16.78 | 145.68 | 2012 | Tropical | 2133 | 2203 | 29 | 29.2 | JR2012 | M | 245 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 502 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 503 | y | y | y |
| *Bactrocera strigifinis* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 504 | y | y | y |
| *Bactrocera tryoni* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 473 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 474 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 475 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 476 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.78 | 145.68 | 2013 | Tropical | 2133 | 1067 | 29 | 29.4 | JR2013 | M | 477 | n | n | y |
| *Dacus axanus* | Cairns | 16.84 | 145.74 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 575 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 481 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 482 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera frauenfeldi* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 488 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 489 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 490 | y | y | KC581378 |
| *Bactrocera frauenfeldi* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 491 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 492 | y | y | KC581378 |
| *Bactrocera neohumeralis* | Cairns | 16.88 | 145.72 | 2012 | Tropical | 2014 | 2003 | 29 | 29.2 | JR2012 | M | 220 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.88 | 145.72 | 2012 | Tropical | 2014 | 2003 | 29 | 29.2 | JR2012 | M | 221 | y | y | KC775787 |
| *Bactrocera neohumeralis* | Cairns | 16.88 | 145.72 | 2012 | Tropical | 2014 | 2003 | 29 | 29.2 | JR2012 | M | 246 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.88 | 145.72 | 2012 | Tropical | 2014 | 2003 | 29 | 29.2 | JR2012 | M | 247 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.88 | 145.72 | 2012 | Tropical | 2014 | 2003 | 29 | 29.2 | JR2012 | M | 248 | y | y | KC775788 |
| *Bactrocera neohumeralis* | Cairns | 16.88 | 145.72 | 2012 | Tropical | 2014 | 2003 | 29 | 29.2 | JR2012 | M | 249 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 505 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 506 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 507 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 508 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 468 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 469 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 470 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 471 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.88 | 145.72 | 2013 | Tropical | 2014 | 1067 | 29 | 29.4 | JR2013 | M | 472 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.9 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 156 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.9 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 157 | y | y | KC581372 |
| *Bactrocera frauenfeldi* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 158 | y | n | KC581379 |
| *Bactrocera manskii* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 126 | n | n | y |
| *Bactrocera manskii* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 127 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 154 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 155 | n | n | y |
| *Bactrocera peninsularis* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 153 | n | y | y |
| *Bactrocera tryoni* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 151 | n | n | y |
| *Bactrocera tryoni* | Cairns | 16.90 | 145.74 | 2001 | Tropical | 2014 | 1668 | 29 | 29.4 | 01/197.1 | M | 152 | n | n | KC581398 |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera fallacis* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.5 | M | 132 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.3 | M | 133 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.3 | M | 134 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.3 | M | 135 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.3 | M | 136 | y | y | KC581378 |
| *Bactrocera jarvisi* | Cairns | 16.92 | 145.78 | 2012 | Tropical | 2272 | 2119 | 29 | 29.2 | JR2012 | M | 255 | n | n | y |
| *Bactrocera jarvisi* | Cairns | 16.92 | 145.78 | 2012 | Tropical | 2272 | 2119 | 29 | 29.2 | JR2012 | M | 256 | n | n | y |
| *Bactrocera jarvisi* | Cairns | 16.92 | 145.78 | 2012 | Tropical | 2272 | 2119 | 29 | 29.2 | JR2012 | M | 257 | n | n | y |
| *Bactrocera jarvisi* | Cairns | 16.92 | 145.78 | 2012 | Tropical | 2272 | 2119 | 29 | 29.2 | JR2012 | M | 258 | n | n | y |
| *Bactrocera jarvisi* | Cairns | 16.92 | 145.78 | 2012 | Tropical | 2272 | 2119 | 29 | 29.2 | JR2012 | M | 259 | n | n | y |
| *Bactrocera jarvisi* | Cairns | 16.92 | 145.78 | 2012 | Tropical | 2272 | 2119 | 29 | 29.2 | JR2012 | M | 260 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.1 | M | 160 | y | n | KC581391 |
| *Bactrocera neohumeralis* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.1 | M | 161 | n | n | y |
| *Bactrocera strigifinis* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.1 | M | 159 | n | n | KC581395 |
| *Bactrocera tryoni* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.1 | M | 275 | y | y | KC581401 |
| *Bactrocera tryoni* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.1 | M | 276 | y | y | KC581402 |
| *Bactrocera tryoni* | Cairns | 16.92 | 145.78 | 2001 | Tropical | 2272 | 2029 | 29 | 29.4 | 01/185A.1 | M | 277 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.95 | 145.73 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/201 | M | 43 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.95 | 145.73 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/201 | M | 44 | n | n | y |
| *Bactrocera fallacis* | Cairns | 16.95 | 145.73 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/201 | M | 17 | n | n | KC581381 |
| *Bactrocera bryoniae* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 147 | n | n | y |
| *Bactrocera bryoniae* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 148 | n | n | y |
| *Bactrocera frauenfeldi* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 150 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 34 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 35 | y | y | KC581390 |
| *Bactrocera neohumeralis* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 107 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 108 | n | n | y |
| *Bactrocera neohumeralis* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 109 | y | y | KC581385 |
| *Bactrocera quadrata* | Cairns | 16.96 | 145.72 | 1998 | Tropical | 2272 | 2084 | 29 | 30 | 98/182b | M | 149 | n | n | KC693011 |
| *Bactrocera bryoniae* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 146 | y | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 104 | n | n | y |
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 105 | n | n | y |
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 106 | n | n | KC581384 |
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk1 | M | 142 | n | n | KC775791 |
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk1 | M | 143 | n | n | KC775792 |
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 144 | n | n | y |
| *Bactrocera neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 145 | n | n | y |
| *Bactrocera strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 49 | n | n | y |
| *Bactrocera strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 50 | n | n | y |
| *Bactrocera strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 138 | y | n | KC581394 |
| *Bactrocera strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 139 | n | n | y |
| *Bactrocera strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 140 | n | n | y |
| *Bactrocera strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182a | M | 141 | n | n | y |
| *Bactrocera tryoni* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk1 | M | 110 | n | n | y |
| *Bactrocera tryoni* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk1 | M | 111 | n | n | y |
| *Bactrocera tryoni* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk1 | M | 112 | n | n | y |
| *Bactrocera tryoni* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 113 | n | n | y |
| *Bactrocera tryoni* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 114 | n | n | y |
| *Bactrocera tryoni* | Gordonvale | 17.07 | 145.77 | 1998 | Tropical | 1878 | 1931 | 29 | 30 | 98/182aWk2 | M | 115 | n | n | y |
| *Bactrocera cacuminata* | Atherton | 17.26 | 145.49 | 2013 | Tropical | 1387 | 1116 | 27.3 | 27.7 | MR2013 | M | 519 | n | n | y |
| *Bactrocera cacuminata* | Atherton | 17.26 | 145.49 | 2013 | Tropical | 1387 | 1116 | 27.3 | 27.7 | MR2013 | M | 520 | n | n | y |
| *Bactrocera cacuminata* | Atherton | 17.26 | 145.49 | 2013 | Tropical | 1387 | 1116 | 27.3 | 27.7 | MR2013 | F | 521 | n | n | y |
| *Bactrocera cacuminata* | Atherton | 17.26 | 145.49 | 2013 | Tropical | 1387 | 1116 | 27.3 | 27.7 | MR2013 | F | 522 | n | n | y |
| *Bactrocera cacuminata* | Atherton | 17.26 | 145.49 | 2013 | Tropical | 1387 | 1116 | 27.3 | 27.7 | MR2013 | F | 523 | n | n | y |
| *Dacus aequalis* | Atherton | 17.26 | 145.49 | 2000 | Tropical | 1387 | 1875 | 27.3 | 26.7 | 00/243.1 | M | 387 | n | n | y |
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | M | 524 | n | n | y |
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | M | 525 | n | n | y |
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | M | 526 | n | n | y |
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | F | 527 | n | n | y |
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | F | 528 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | F | 529 | n | n | y |
| *Bactrocera cacuminata* | Herberton | 17.38 | 145.39 | 2013 | Tropical | 1157 | 794 | 25.4 | 27.7 | MR2013 | F | 530 | n | n | y |
| *Bactrocera tryoni* | Karumba | 17.49 | 140.88 | 1998 | Tropical | 891 | 1463 | 33.3 | 32.7 | 98/A934 | M | 305 | n | n | y |
| *Bactrocera tryoni* | Karumba | 17.49 | 140.88 | 1998 | Tropical | 891 | 1463 | 33.3 | 32.7 | 98/A934 | M | 306 | n | n | y |
| *Bactrocera tryoni* | Karumba | 17.49 | 140.88 | 1998 | Tropical | 891 | 1463 | 33.3 | 32.7 | 98/A934 | M | 307 | n | n | y |
| *Bactrocera tryoni* | Karumba | 17.49 | 140.88 | 1998 | Tropical | 891 | 1463 | 33.3 | 32.7 | 98/A934 | M | 308 | n | n | y |
| *Bactrocera tryoni* | Karumba | 17.49 | 140.88 | 1998 | Tropical | 891 | 1463 | 33.3 | 32.7 | 98/A934 | M | 309 | n | n | y |
| *Bactrocera tryoni* | Karumba | 17.49 | 140.88 | 1998 | Tropical | 891 | 1463 | 33.3 | 32.7 | 98/A934 | M | 310 | n | n | y |
| *Dacus newmani* | Karumba | 17.49 | 140.88 | 1997 | Tropical | 891 | 874 | 33.3 | 32.1 | A929 | M | 69 | n | n | KC581411 |
| *Bactrocera bryoniae* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 547 | n | n | y |
| *Bactrocera bryoniae* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 548 | n | n | y |
| *Bactrocera bryoniae* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 549 | n | n | y |
| *Bactrocera bryoniae* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 550 | n | n | y |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.59 | 146.07 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 238 | y | y | KC775789 |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.59 | 146.07 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 239 | n | n | y |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.59 | 146.07 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 250 | n | n | y |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.59 | 146.07 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 251 | n | n | y |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.59 | 146.07 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 252 | n | n | y |
| *Bactrocera peninsularis* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 532 | n | y | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 463 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 464 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 465 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 466 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 467 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 531 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 533 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.59 | 146.07 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 534 | n | n | y |
| *Bactrocera bryoniae* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 551 | n | n | y |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.60 | 146.13 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 240 | y | y | KC775790 |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.60 | 146.13 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 241 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.60 | 146.13 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 253 | n | n | y |
| *Bactrocera neohumeralis* | Mourilyan Harbour | 17.60 | 146.13 | 2012 | Tropical | 3575 | 3366 | 27.9 | 28 | JR2012 | M | 254 | n | n | y |
| *Bactrocera strigifinis* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 501 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 458 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 459 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 460 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 461 | n | n | y |
| *Bactrocera tryoni* | Mourilyan Harbour | 17.60 | 146.13 | 2013 | Tropical | 3575 | 3082 | 27.9 | 27.9 | JR2013 | M | 462 | n | n | y |
| *Bactrocera manskii* | Tully | 17.93 | 145.92 | 2001 | Tropical | 4127 | 3144 | 28.7 | 28.5 | 01/181B.1 | M | 124 | n | n | y |
| *Bactrocera manskii* | Tully | 17.93 | 145.92 | 2001 | Tropical | 4127 | 3144 | 28.7 | 28.5 | 01/181B.1 | M | 125 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.23 | 146.67 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 224 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.23 | 146.67 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 225 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.23 | 146.67 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 383 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.23 | 146.67 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 384 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.23 | 146.67 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 385 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.23 | 146.67 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 386 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.23 | 146.67 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 448 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.23 | 146.67 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 449 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.23 | 146.67 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 450 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.23 | 146.67 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 451 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.23 | 146.67 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 452 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 222 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 223 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 379 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 380 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 381 | n | n | y |
| *Bactrocera neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | Tropical | 1087 | 1314 | 28.9 | 28.7 | JR2012 | M | 382 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.36 | 146.84 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 453 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.36 | 146.84 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 454 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.36 | 146.84 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 455 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera tryoni* | Townsville | 19.36 | 146.84 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 456 | n | n | y |
| *Bactrocera tryoni* | Townsville | 19.36 | 146.84 | 2013 | Tropical | 1087 | 655 | 28.9 | 29 | JR2013 | M | 457 | n | n | y |
| *Bactrocera tryoni* | Charters Towers | 20.08 | 146.26 | 1998 | Subtropical | 684 | 1197 | 30.3 | 30.1 | 98/178a | M | 13 | n | n | y |
| *Bactrocera tryoni* | Charters Towers | 20.08 | 146.26 | 1998 | Subtropical | 684 | 1197 | 30.3 | 30.1 | 98/178a | M | 14 | n | n | y |
| *Dacus aequalis* | Charters Towers | 20.08 | 146.26 | 2000 | Subtropical | 684 | | 30.3 | | 00/178A.1 | M | 388 | n | n | y |
| *Bactrocera chorista* | Calen | 20.90 | 148.77 | 2001 | Subtropical | 1898 | 1191 | 28.7 | 29.2 | 01/169A.1 | M | 130 | n | n | y |
| *Bactrocera chorista* | Calen | 20.90 | 148.77 | 2001 | Subtropical | 1898 | 1191 | 28.7 | 29.2 | 01/169A.1 | M | 131 | n | n | y |
| *Bactrocera neohumeralis* | Calen | 20.90 | 148.77 | 1998 | Subtropical | 1898 | 1038 | 28.7 | 28.7 | 98/169 | M | 32 | n | n | y |
| *Bactrocera neohumeralis* | Calen | 20.90 | 148.77 | 1998 | Subtropical | 1898 | 1038 | 28.7 | 28.7 | 98/169 | M | 33 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.08 | 149.19 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 230 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.08 | 149.19 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 231 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.08 | 149.19 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 342 | y | y | y |
| *Bactrocera neohumeralis* | Mackay | 21.08 | 149.19 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 343 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.08 | 149.19 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 443 | y | y | y |
| *Bactrocera tryoni* | Mackay | 21.08 | 149.19 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 444 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.08 | 149.19 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 445 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.08 | 149.19 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 446 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.08 | 149.19 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 447 | n | n | y |
| *Bactrocera bryoniae* | Mackay | 21.11 | 149.10 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 552 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 232 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 233 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 344 | n | n | y |
| *Bactrocera neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 345 | y | y | y |
| *Bactrocera neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 346 | y | y | y |
| *Bactrocera neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | Subtropical | 1886 | 1840 | 27.1 | 26.9 | JR2012 | M | 347 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.11 | 149.10 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 438 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.11 | 149.10 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 439 | y | y | y |
| *Bactrocera tryoni* | Mackay | 21.11 | 149.10 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 440 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.11 | 149.10 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 441 | n | n | y |
| *Bactrocera tryoni* | Mackay | 21.11 | 149.10 | 2013 | Subtropical | 1886 | 1928 | 27.1 | 26.4 | JR2013 | M | 442 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera tryoni* | Emerald | 23.53 | 148.18 | 1998 | Subtropical | 561 | 982 | 29.7 | 29.4 | 98/162a | M | 11 | n | n | y |
| *Bactrocera tryoni* | Emerald | 23.53 | 148.18 | 1998 | Subtropical | 561 | 982 | 29.7 | 29.4 | 98/162a | M | 12 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/21.10 | M | 120 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/21.10 | M | 121 | n | n | KC581397 |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 122 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 123 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 162 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 163 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 164 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 165 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 166 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 167 | n | n | KC581399 |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.12 | M | 169 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.12 | M | 170 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.12 | M | 171 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.9 | M | 173 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.9 | M | 174 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 179 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 180 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 181 | n | n | y |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 182 | n | n | KC581400 |
| *Bactrocera tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 183 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.10 | M | 168 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22A.12 | M | 172 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.9 | M | 175 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.9 | M | 176 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.9 | M | 177 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 184 | n | n | y |
| *Dacus newmani* | Alice Springs | 23.70 | 133.88 | 1998 | Grasslands | 285 | 295 | 28.7 | 29.8 | 98/22B.10 | M | 185 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.84 | 151.26 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 228 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Gladstone | 23.84 | 151.26 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 229 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.84 | 151.26 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 358 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.84 | 151.26 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 359 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.84 | 151.26 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 360 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.84 | 151.26 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 433 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.84 | 151.26 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 434 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.84 | 151.26 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 435 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.84 | 151.26 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 436 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.84 | 151.26 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 437 | n | n | y |
| *Bactrocera bryoniae* | Gladstone | 23.88 | 151.24 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 478 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.88 | 151.24 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 226 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.88 | 151.24 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 227 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.88 | 151.24 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 355 | y | y | y |
| *Bactrocera neohumeralis* | Gladstone | 23.88 | 151.24 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 356 | n | n | y |
| *Bactrocera neohumeralis* | Gladstone | 23.88 | 151.24 | 2012 | Subtropical | 886 | 764 | 27.7 | 27.8 | JR2012 | M | 357 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.88 | 151.24 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 428 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.88 | 151.24 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 429 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.88 | 151.24 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 430 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.88 | 151.24 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 431 | n | n | y |
| *Bactrocera tryoni* | Gladstone | 23.88 | 151.24 | 2013 | Subtropical | 886 | 1431 | 27.7 | 27.8 | JR2013 | M | 432 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.77 | 152.38 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 236 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.77 | 152.38 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 237 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.77 | 152.38 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 348 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.77 | 152.38 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 350 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.77 | 152.38 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 351 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.77 | 152.38 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 423 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.77 | 152.38 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 424 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.77 | 152.38 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 425 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.77 | 152.38 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 426 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.77 | 152.38 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 427 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Bundaberg | 24.89 | 152.37 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 234 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.89 | 152.37 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 235 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.89 | 152.37 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 352 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.89 | 152.37 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 353 | n | n | y |
| *Bactrocera neohumeralis* | Bundaberg | 24.89 | 152.37 | 2012 | Subtropical | 1032 | 1059 | 26.6 | 26.7 | JR2012 | M | 354 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.89 | 152.37 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 418 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.89 | 152.37 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 419 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.89 | 152.37 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 420 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.89 | 152.37 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 421 | n | n | y |
| *Bactrocera tryoni* | Bundaberg | 24.89 | 152.37 | 2013 | Subtropical | 1032 | 1190 | 26.6 | 27 | JR2013 | M | 422 | n | n | y |
| *Bactrocera tryoni* | Kingaroy | 26.54 | 151.84 | 1998 | Subtropical | 774 | 1016 | 25.6 | 27.1 | 98/153a | M | 9 | n | n | y |
| *Bactrocera tryoni* | Kingaroy | 26.54 | 151.84 | 1998 | Subtropical | 774 | 1016 | 25.6 | 27.1 | 98/153a | M | 10 | n | n | y |
| *Bactrocera tryoni* | Woombye | 26.67 | 152.98 | 1998 | Subtropical | 1689 | 1363 | 24.2 | 25.4 | 98/151a | M | 7 | n | n | y |
| *Bactrocera tryoni* | Woombye | 26.67 | 152.98 | 1998 | Subtropical | 1689 | 1363 | 24.2 | 25.4 | 98/151a | M | 8 | n | n | y |
| *Bactrocera tryoni* | Nanango | 26.67 | 151.99 | 1998 | Subtropical | 790 | 884 | 25.3 | 27.1 | 98/154.1 | M | 290 | n | n | y |
| *Bactrocera tryoni* | Nanango | 26.67 | 151.99 | 1998 | Subtropical | 790 | 884 | 25.3 | 27.1 | 98/154.1 | M | 291 | n | n | y |
| *Bactrocera tryoni* | Nanango | 26.67 | 151.99 | 1998 | Subtropical | 790 | 884 | 25.3 | 27.1 | 98/154.1 | M | 292 | n | n | y |
| *Bactrocera neohumeralis* | Burpengary | 27.15 | 152.97 | 2001 | Subtropical | 1264 | 979 | 25.9 | 26.8 | 01/145.1 | M | 281 | n | n | y |
| *Bactrocera neohumeralis* | Burpengary | 27.15 | 152.97 | 2001 | Subtropical | 1264 | 979 | 25.9 | 26.8 | 01/145.1 | M | 282 | n | n | y |
| *Bactrocera tryoni* | Burpengary | 27.15 | 152.97 | 2001 | Subtropical | 1264 | 979 | 25.9 | 26.8 | 01/145.1 | M | 278 | n | n | y |
| *Bactrocera tryoni* | Burpengary | 27.15 | 152.97 | 2001 | Subtropical | 1264 | 979 | 25.9 | 26.8 | 01/145.1 | M | 279 | n | n | y |
| *Bactrocera tryoni* | Burpengary | 27.15 | 152.97 | 2001 | Subtropical | 1264 | 979 | 25.9 | 26.8 | 01/145.1 | M | 280 | n | n | y |
| *Bactrocera tryoni* | Dalby | 27.18 | 151.27 | 1998 | Subtropical | 626 | 794 | 26.7 | 26.4 | 98/135a | M | 5 | n | n | y |
| *Bactrocera tryoni* | Dalby | 27.18 | 151.27 | 1998 | Subtropical | 626 | 794 | 26.7 | 26.4 | 98/135a | M | 6 | n | n | y |
| *Bactrocera bryoniae* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 41 | n | n | y |
| *Bactrocera bryoniae* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 42 | n | n | y |
| *Bactrocera chorista* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 20 | n | n | y |
| *Bactrocera chorista* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 21 | n | n | y |
| *Bactrocera neohumeralis* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 30 | n | n | y |
| *Bactrocera neohumeralis* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 31 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera quadrata* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 372 | n | n | y |
| *Bactrocera quadrata* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 373 | n | n | y |
| *Dacus absonifacies* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147.5 | M | 376 | n | n | y |
| *Dacus aequalis* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147 | M | 53 | n | n | y |
| *Dacus aequalis* | Deception Bay | 27.22 | 153.09 | 1996 | Subtropical | 1182 | 1524 | 25.2 | 25.1 | 147.8 | M | 96 | n | n | y |
| *Dacus aequalis* | Deception Bay | 27.22 | 153.09 | 1996 | Subtropical | 1182 | 1524 | 25.2 | 25.1 | 147.8 | M | 97 | n | n | y |
| *Dacus newmani* | Deception Bay | 27.22 | 153.09 | 1998 | Subtropical | 1182 | 1187 | 25.2 | 25.4 | 98/147.10 | M | 377 | n | n | y |
| *Bactrocera chorista* | Kallangur | 27.25 | 152.99 | 1998 | Subtropical | 1169 | 905 | 25.2 | 25.4 | 98/144.1 | M | 364 | n | n | y |
| *Bactrocera chorista* | Herron Park | 27.33 | 152.85 | 1998 | Subtropical | 1169 | 905 | 25.2 | 25.4 | 98/149 | M | 362 | n | n | y |
| *Bactrocera chorista* | Herron Park | 27.33 | 152.85 | 1998 | Subtropical | 1169 | 905 | 25.2 | 25.4 | 98/149 | M | 363 | n | n | y |
| *Bactrocera tryoni* | Herron Park | 27.33 | 152.85 | 1998 | Subtropical | 1169 | 905 | 25.2 | 25.4 | 98/149 | M | 287 | n | n | y |
| *Bactrocera tryoni* | Herron Park | 27.33 | 152.85 | 1998 | Subtropical | 1169 | 905 | 25.2 | 25.4 | 98/149 | M | 288 | n | n | y |
| *Bactrocera tryoni* | Herron Park | 27.33 | 152.85 | 1998 | Subtropical | 1169 | 905 | 25.2 | 25.4 | 98/149 | M | 289 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 153.10 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 328 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 153.10 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 329 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 153.10 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 330 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 153.10 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 331 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 153.10 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 332 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 153.10 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 333 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 153.10 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 414 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 153.10 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 416 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 153.10 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 417 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 153.10 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 415a | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 153.10 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 415b | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 152.99 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 322 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 152.99 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 323 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 152.99 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 324 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 152.99 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 325 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 152.99 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 326 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (N) | 27.35 | 152.99 | 2012 | Subtropical | 1160 | 1583 | 25.3 | 24.8 | JR2012 | M | 327 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 152.99 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 409 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 152.99 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 410 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 152.99 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 411 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 152.99 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 412 | n | n | y |
| *Bactrocera tryoni* | Brisbane (N) | 27.35 | 152.99 | 2013 | Subtropical | 1160 | 1384 | 25.3 | 24.9 | JR2013 | M | 413 | n | n | y |
| *Bactrocera cacuminata* | Brisbane | 27.51 | 152.94 | 1997 | Subtropical | 1017 | 1132 | 26.2 | 26.1 | 12/Bca4 | M | 393 | n | n | y |
| *Bactrocera cacuminata* | Brisbane | 27.51 | 152.94 | 1997 | Subtropical | 1017 | 1132 | 26.2 | 26.1 | 12/Bca4 | M | 394 | n | n | y |
| *Bactrocera cacuminata* | Brisbane | 27.51 | 152.94 | 1997 | Subtropical | 1017 | 1132 | 26.2 | 26.1 | 12/Bca4 | M | 395 | n | n | y |
| *Bactrocera cacuminata* | Brisbane | 27.51 | 152.94 | 1997 | Subtropical | 1017 | 1132 | 26.2 | 26.1 | 12/Bca4 | M | 396 | n | n | y |
| *Bactrocera cacuminata* | Brisbane | 27.51 | 152.94 | 1997 | Subtropical | 1017 | 1132 | 26.2 | 26.1 | 12/Bca4 | M | 397 | n | n | y |
| *Bactrocera cacuminata* | Brisbane | 27.51 | 152.94 | 1997 | Subtropical | 1017 | 1132 | 26.2 | 26.1 | 12/Bca4 | M | 398 | n | n | y |
| *Bactrocera bryoniae* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 | M | 39 | n | n | y |
| *Bactrocera bryoniae* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/132wk2 | M | 40 | n | n | y |
| *Bactrocera chorista* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 | M | 18 | n | n | KC581375 |
| *Bactrocera chorista* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 | M | 19 | n | n | y |
| *Bactrocera quadrata* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 wk1 | M | 368 | n | n | y |
| *Bactrocera quadrata* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 wk1 | M | 369 | n | n | y |
| *Bactrocera quadrata* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 wk3 | M | 370 | n | n | y |
| *Bactrocera quadrata* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133wk3 | M | 371 | n | n | y |
| *Bactrocera quadrata* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 wk4 | M | 374 | n | n | y |
| *Bactrocera quadrata* | Toowoomba | 27.54 | 151.91 | 1998 | Subtropical | 842 | 821 | 24.3 | 23.1 | 98/133 wk4 | M | 375 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (S) | 27.58 | 153.06 | 1998 | Subtropical | 1081 | 1156 | 26.2 | 26.8 | 98/123Wk1 | M | 26 | n | n | KC581388 |
| *Bactrocera neohumeralis* | Brisbane (S) | 27.58 | 153.06 | 1998 | Subtropical | 1081 | 1156 | 26.2 | 26.8 | 98/123Wk2 | M | 27 | n | n | KC581389 |
| *Bactrocera chorista* | Brisbane (S) | 27.58 | 153.06 | 1998 | Subtropical | 1081 | 1156 | 26.2 | 26.8 | 98/123 | M | 361 | n | n | y |
| *Bactrocera neohumeralis* | North Ipswich | 27.59 | 152.76 | 1998 | Subtropical | 879 | 674 | 27.3 | 27.5 | 98/130 | M | 28 | n | n | y |
| *Bactrocera neohumeralis* | North Ipswich | 27.59 | 152.76 | 1998 | Subtropical | 879 | 674 | 27.3 | 27.5 | 98/130 | M | 29 | n | n | y |
| *Bactrocera tryoni* | Redbank Plains | 27.65 | 152.85 | 1998 | Subtropical | 864 | 674 | 26.7 | 27.5 | 98/129 | M | 3 | n | n | y |
| *Bactrocera tryoni* | Redbank Plains | 27.65 | 152.85 | 1998 | Subtropical | 864 | 674 | 26.7 | 27.5 | 98/129 | M | 4 | n | n | y |
| *Bactrocera bryoniae* | Brisbane (SE) | 27.66 | 152.98 | 2013 | Subtropical | 1017 | 1096 | 26.2 | 26.6 | JR2013 | M | 553 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.66 | 152.98 | 2012 | Subtropical | 1017 | 1132 | 26.2 | 26.6 | JR2012 | M | 338 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.66 | 152.98 | 2012 | Subtropical | 1017 | 1132 | 26.2 | 26.6 | JR2012 | M | 339 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.66 | 152.98 | 2012 | Subtropical | 1017 | 1132 | 26.2 | 26.6 | JR2012 | M | 340 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.66 | 152.98 | 2012 | Subtropical | 1017 | 1132 | 26.2 | 26.6 | JR2012 | M | 341 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.66 | 152.98 | 2013 | Subtropical | 1017 | 1096 | 26.2 | 26.6 | JR2013 | M | 399 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.66 | 152.98 | 2013 | Subtropical | 1017 | 1096 | 26.2 | 26.6 | JR2013 | M | 400 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.66 | 152.98 | 2013 | Subtropical | 1017 | 1096 | 26.2 | 26.6 | JR2013 | M | 401 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.66 | 152.98 | 2013 | Subtropical | 1017 | 1096 | 26.2 | 26.6 | JR2013 | M | 402 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.66 | 152.98 | 2013 | Subtropical | 1017 | 1096 | 26.2 | 26.6 | JR2013 | M | 403 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.70 | 153.20 | 2012 | Subtropical | 1090 | 1385 | 25.9 | 25.9 | JR2012 | M | 334 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.70 | 153.20 | 2012 | Subtropical | 1090 | 1385 | 25.9 | 25.9 | JR2012 | M | 335 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.70 | 153.20 | 2012 | Subtropical | 1090 | 1385 | 25.9 | 25.9 | JR2012 | M | 336 | n | n | y |
| *Bactrocera neohumeralis* | Brisbane (SE) | 27.70 | 153.20 | 2012 | Subtropical | 1090 | 1385 | 25.9 | 25.9 | JR2012 | M | 337 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.70 | 153.20 | 2013 | Subtropical | 1090 | 1416 | 26.2 | 26 | JR2013 | M | 404 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.70 | 153.20 | 2013 | Subtropical | 1090 | 1416 | 26.2 | 26 | JR2013 | M | 405 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.70 | 153.20 | 2013 | Subtropical | 1090 | 1416 | 26.2 | 26 | JR2013 | M | 406 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.70 | 153.20 | 2013 | Subtropical | 1090 | 1416 | 26.2 | 26 | JR2013 | M | 407 | n | n | y |
| *Bactrocera tryoni* | Brisbane (SE) | 27.70 | 153.20 | 2013 | Subtropical | 1090 | 1416 | 26.2 | 26 | JR2013 | M | 408 | n | n | y |
| *Bactrocera bryoniae* | Murwillumbah | 28.33 | 153.40 | 1997 | Subtropical | 1610 | 1336 | 25.7 | 25.7 | 97/70.1 | M | 392 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk1 | M | 24 | n | n | KC581387 |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk2 | M | 25 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk1 | M | 311 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk1 | M | 312 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk1 | M | 313 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk1 | M | 314 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk2 | M | 315 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk2 | M | 316 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk2 | M | 317 | n | n | y |
| *Bactrocera neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | Temperate | 1610 | 1150 | 26.8 | 26.2 | 98/68aWk2 | M | 318 | n | n | y |
| *Bactrocera neohumeralis* | Coffs Harbour | 30.23 | 153.15 | 1998 | Temperate | 1473 | 995 | 23.3 | 23.7 | 98/64Wk1 | M | 22 | n | n | y |
| *Bactrocera neohumeralis* | Coffs Harbour | 30.23 | 153.15 | 1998 | Temperate | 1473 | 995 | 23.3 | 23.7 | 98/64Wk2 | M | 23 | n | n | KC581386 |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 1996 | Temperate | 1473 | 1900 | 23.3 | 23.3 | 96/64.2 | M | 293 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 1996 | Temperate | 1473 | 1900 | 23.3 | 23.3 | 96/64.2 | M | 294 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 1996 | Temperate | 1473 | 1900 | 23.3 | 23.3 | 96/64.2 | M | 295 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 2001 | Temperate | 1473 | 982 | 23.3 | 23.5 | 01/64.3 | M | 577 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 2001 | Temperate | 1473 | 982 | 23.3 | 23.5 | 01/64.3 | M | 579 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 2001 | Temperate | 1473 | 982 | 23.3 | 23.5 | 01/64.3 | M | 580 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 2001 | Temperate | 1473 | 982 | 23.3 | 23.5 | 01/64.3 | M | 581 | n | n | y |
| *Bactrocera tryoni* | Coffs Harbour | 30.23 | 153.15 | 2001 | Temperate | 1473 | 982 | 23.3 | 23.5 | 01/64.3 | M | 582 | n | n | y |
| *Bactrocera tryoni* | Tamworth | 31.09 | 150.93 | 1997 | Temperate | 628 | 576 | 24.6 | 24.8 | 97/53.1 | M | 302 | n | n | y |
| *Bactrocera tryoni* | Tamworth | 31.09 | 150.93 | 1997 | Temperate | 628 | 576 | 24.6 | 24.8 | 97/53.1 | M | 303 | n | n | y |
| *Bactrocera tryoni* | Tamworth | 31.09 | 150.93 | 1997 | Temperate | 628 | 576 | 24.6 | 24.8 | 97/53.1 | M | 304 | n | n | y |
| *Bactrocera tryoni* | Condobolin | 33.07 | 147.15 | 1998 | Temperate | 459 | 566 | 24.4 | 24.1 | 98/109b | M | 284 | n | n | y |
| *Bactrocera tryoni* | Condobolin | 33.07 | 147.15 | 1998 | Temperate | 459 | 566 | 24.4 | 24.1 | 98/109b | M | 285 | n | n | y |
| *Bactrocera tryoni* | Condobolin | 33.07 | 147.15 | 1998 | Temperate | 459 | 566 | 24.4 | 24.1 | 98/109b | M | 286 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C1 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C10 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C12 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C13 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C15 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C2 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C3 | n | n | y |
| *Bactrocera tryoni* | Richmond | 33.35 | 150.45 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | C4 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.35 | 150.45 | 2010 | Temperate | 800 | 792 | 24 | 23.7 | MR2010 | F | 77 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.35 | 150.45 | 2010 | Temperate | 800 | 792 | 24 | 23.7 | MR2010 | F | 78 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.35 | 150.45 | 2010 | Temperate | 800 | 792 | 24 | 23.7 | MR2010 | F | 79 | n | n | KC581405 |
| *Bactrocera tryoni* | Forbes | 33.38 | 148.00 | 1998 | Temperate | 503 | 580 | 24.5 | 23.7 | 98/1086 | M | 1 | n | n | y |
| *Bactrocera tryoni* | Forbes | 33.38 | 148.00 | 1998 | Temperate | 503 | 580 | 24.5 | 23.7 | 98/1086 | M | 2 | n | n | y |
| *Dacus aequalis* | Narara | 33.40 | 151.35 | 1998 | Temperate | 1329 | 1604 | 23 | 23.2 | 98/45 | M | 51 | n | n | y |
| *Dacus aequalis* | Narara | 33.40 | 151.35 | 1998 | Temperate | 1329 | 1604 | 23 | 23.2 | 98/45 | M | 52 | n | n | KC581408 |
| *Bactrocera tryoni* | Gosford | 33.42 | 151.34 | 1997 | Temperate | 1356 | 1093 | 21.9 | 22.3 | 97/45.1 | M | 299 | n | n | y |

| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | Climate zone | Mean rainfall (mm) | Year rainfall (mm) | Mean maximum temp (°C) | Year maximum temp (°C) | Stock no. | Gender M/F | ID No. | *wsp* | 16S | *COI* sequence accession no. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Bactrocera tryoni* | Gosford | 33.42 | 151.34 | 1997 | Temperate | 1356 | 1093 | 21.9 | 22.3 | 97/45.1 | M | 300 | n | n | y |
| *Bactrocera tryoni* | Gosford | 33.42 | 151.34 | 1997 | Temperate | 1356 | 1093 | 21.9 | 22.3 | 97/45.1 | M | 301 | n | n | y |
| *Dacus aequalis* | Gosford | 33.42 | 151.34 | 1996 | Temperate | 1356 | 1162 | 21.9 | 21.5 | 96/45.2 | M | 98 | n | n | y |
| *Dacus aequalis* | Gosford | 33.42 | 151.34 | 1996 | Temperate | 1356 | 1162 | 21.9 | 21.5 | 96/45.2 | M | 99 | n | n | y |
| *Dacus absonifacies* | Kurrajong | 33.57 | 150.67 | 1996 | Temperate | 1251 | 1055 | 24 | 23.6 | 96/87 | M | 100 | n | n | y |
| *Dacus absonifacies* | Kurrajong | 33.57 | 150.67 | 1996 | Temperate | 1251 | 1055 | 24 | 23.6 | 96/87 | M | 101 | n | n | y |
| *Dacus absonifacies* | Kurrajong | 33.57 | 150.67 | 1996 | Temperate | 1251 | 1055 | 24 | 23.6 | 96/87 | M | 102 | n | n | y |
| *Dacus aequalis* | Kurrajong | 33.57 | 150.67 | 1996 | Temperate | 1251 | 1055 | 24 | 23.6 | 96/87 | M | 94 | n | n | y |
| *Dacus aequalis* | Kurrajong | 33.57 | 150.67 | 1996 | Temperate | 1251 | 1055 | 24 | 23.6 | 96/87 | M | 95 | n | n | y |
| *Dirioxa pornia* | Kurrajong | 33.57 | 150.67 | 1996 | Temperate | 1251 | 1055 | 24 | 23.6 | 96/87 | M | 103 | n | n | y |
| *Bactrocera cacuminata* | Richmond | 33.60 | 150.75 | 2011 | Temperate | 800 | 725 | 24 | 23.4 | MR2011 | M | 186 | n | n | KC581374 |
| *Bactrocera cacuminata* | Richmond | 33.60 | 150.75 | 2011 | Temperate | 800 | 725 | 24 | 23.4 | MR2011 | M | 187 | n | n | y |
| *Bactrocera cacuminata* | Richmond | 33.60 | 150.75 | 2011 | Temperate | 800 | 725 | 24 | 23.4 | MR2011 | M | 188 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A1 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A10 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A19 | n | n | KC581406 |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A2 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A20 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A3 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A4 | n | n | y |
| *Dirioxa pornia* | Richmond | 33.60 | 150.75 | 2012 | Temperate | 800 | 878 | 24 | 23.1 | MR2012 | F | A6 | n | n | y |
| *Bactrocera tryoni* | Gerringong | 34.75 | 150.83 | 1996 | Temperate | 1248 | 1103 | 21.5 | 21.3 | 96/73 | M | 296 | n | n | y |
| *Bactrocera tryoni* | Gerringong | 34.75 | 150.83 | 1996 | Temperate | 1248 | 1103 | 21.5 | 21.3 | 96/73 | M | 297 | n | n | y |
| *Bactrocera tryoni* | Gerringong | 34.75 | 150.83 | 1996 | Temperate | 1248 | 1103 | 21.5 | 21.3 | 96/73 | M | 298 | n | n | y |
| *Dacus aequalis* | Geringong | 34.75 | 150.83 | 1996 | Temperate | 1248 | 1103 | 21.5 | 21.3 | 96/73 | M | 92 | n | n | y |
| *Dacus aequalis* | Geringong | 34.75 | 150.83 | 1996 | Temperate | 1248 | 1103 | 21.5 | 21.3 | 96/73 | M | 93 | n | n | y |
| *Dacus absonifacies* | Dalmeny | 36.17 | 150.13 | 1998 | Temperate | 970 | 1064 | 20.4 | 20.8 | 98/74a | M | 54 | n | n | y |
| *Dacus absonifacies* | Dalmeny | 36.17 | 150.13 | 1998 | Temperate | 970 | 1064 | 20.4 | 20.8 | 98/74a | M | 55 | n | n | KC581407 |

**Table A.2** Laboratory strains screened for *Wolbachia* using *wsp* and 16S rRNA loci.

| Species | Stock line | Laboratory location* | Sampling year | Individual name | 16S | *wsp* | COI accession no. |
|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn1 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn2 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn3 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn4 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn5 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn6 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn7 | n | n | |
| *Bactrocera neohumeralis* | Bneo | Richmond, UWS | 2012 | Bn8 | n | n | |
| | | | | | | | |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac1 | n | n | KC581374 |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac2 | n | n | KC581374 |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac3 | n | n | |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac4 | n | n | |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac5 | n | n | |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac6 | n | n | |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac7 | n | n | |
| *Bactrocera cacuminata* | Bcac | Richmond, UWS | 2012 | Bcac8 | n | n | |
| | | | | | | | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC1 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC2 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC3 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC4 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC5 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC6 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC7 | n | n | |
| *Bactrocera tryoni* | BtHAC | Richmond, UWS | 2012 | HAC8 | n | n | |
| | | | | | | | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS1 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS2 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS3 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS4 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS5 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS6 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS7 | n | n | |
| *Bactrocera tryoni* | BtGOS | Richmond, UWS | 2012 | GOS8 | n | n | |
| | | | | | | | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj1 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj2 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj3 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj4 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj5 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj6 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj7 | n | n | |
| *Bactrocera jarvisi* | Bjar | Richmond, UWS | 2012 | Bj8 | n | n | |

| Species | Stock line | Laboratory location* | Sampling year | Individual name | 16S | *wsp* | COI accession no. |
|---|---|---|---|---|---|---|---|
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-1 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-2 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-3 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-4 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-5 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-6 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-7 | n | n | |
| *Bactrocera neohumeralis* | Bneo09 | Cairns, QDAFF | 2012 | Bn9-8 | n | n | |
| | | | | | | | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-1 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-2 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-3 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-4 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-5 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-6 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-7 | n | n | |
| *Bactrocera neohumeralis* | Bneo12 | Cairns, QDAFF | 2012 | Bn12-8 | n | n | |
| | | | | | | | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF1 | n | n | KC581412 |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF2 | n | n | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF3 | n | n | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF4 | n | n | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF5 | n | n | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF6 | n | n | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF7 | n | n | |
| *Ceratitis capitata* | MedFly | Perth, DAFWA | 2011 | MF8 | n | n | |

*UWS (University of Western Sydney), QDAFF (Queensland Department of Agriculture, Fisheries and Forestry), DAFWA (Department of Agriculture and Food, Western Australia)

**Table A.3** PCR conditions and primer sequences used in this research.

| | *wsp* | 16S rDNA | *COI* and Sequencing | DIG-labelled *wsp* probe |
|---|---|---|---|---|
| **PCR protocol** | | | | |
| Total volume | 10µL / 20µL | 10µL | 20µL | 50µL |
| GoTaq Reaction Buffer | 1X | 1X | 1X | 1X |
| $MgCl_2$ | 4mM | 2mM | 2.5mM | 5mM |
| dNTP | 150µM each | 125µM each | 200µM each | 2.5µL (DIG-dNTP labelling mixture) |
| Primer 1 | 0.8µM | 0.63µM | 0.4µM | 0.6µM |
| Primer 2 | 0.8µM | 0.63µM | 0.4µM | 0.6µM |
| GoTaq DNA polymerase | 0.4u / 0.8u | 0.5u | 0.5u | 1u |
| DNA Template | 1µL / 2µL | 1µL | 2µL | 2µL |
| | | | | |
| **Thermal Cycling** | | | | |
| Denaturation | 94°C for 2min | 94°C for 5min | 94°C for 5min | 94°C for 2min |
| | 94°C for 10s, 65°C for 30s, 68°C for 1min (x10 cycles) | 94°C for 30s, 54°C for 30s, 72°C for 1min (x35 cycles) | 94°C for 30s, 55°C for 30s, 72°C for 1min (x35 cycles) | 94°C for 10s, 65°C for 30s, 68°C for 1min (x10 cycles) |
| | 94°C for 10s, 65°C for 30s, 68°C for 1min (x25 cycles, including an incremental increase of 20s at 68°C) | 72°C for 10min | 72°C for 10min | 94°C for 10s, 65°C for 30s, 68°C for 1min (x25 cycles, including an incremental increase of 20s at 68°C) |
| | | | | |
| **Primers** | | | | |
| *wsp* | | | | |
| Wsp-F | TGGTCCAATAAGTGATGAAGAAACTAGCTA | Jeyaprakash and Hoy, 2000 | | |
| Wsp-R | AAAAATTAAACGCTACTCCAGCTTCTGCAC | Jeyaprakash and Hoy, 2000 | | |
| 16S rDNA | | | | |
| wspecF | CATACCTATTCGAAGGGATAG | Werren and Windsor, 2000 | | |
| wspecR | AGCTTCGAGTGAAACCAATTC | Werren and Windsor, 2000 | | |
| *COI* | | | | |
| Dick (C1-J-2441) | CCTACAGGAATTAAAATTTTTAGATGATTA | Simon et al. 1994 | | |
| Pat (TL2-N-3014) | TCCATTGCACTAATCTGCCATATTA | Simon et al. 1994 | | |
| Sequencing | | | | |
| T7 Promoter | TAATACGACTCACTATAGGG | | | |
| SP6 | ATTTAGGTGACACTATAG | | | |

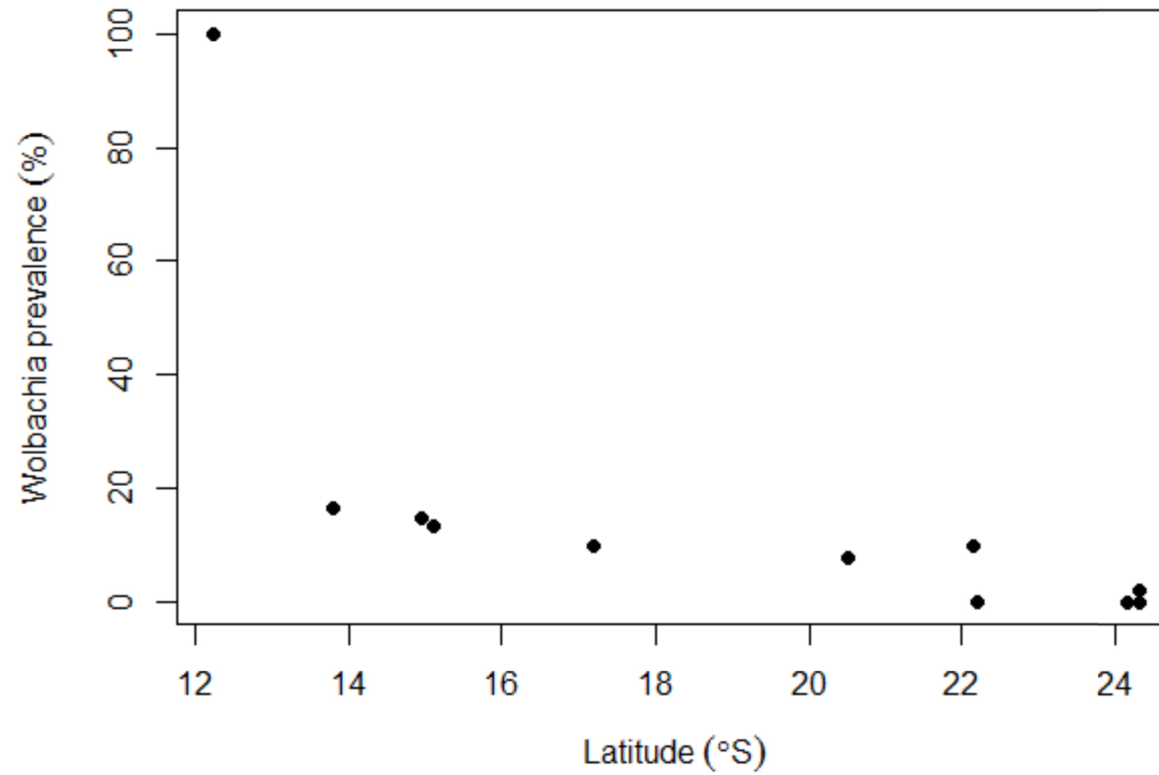



**Figure A.1** Relationship between prevalence of *Wolbachia* infection in tephritid species and the midpoint of the latitudinal distribution, based on Hancock *et al*. (2000) and Royer and Hancock (2012) ($R^2$ = 0.46, $F_{1,9}$ = 9.33; $p < 0.05$). At least ten individuals were tested for uninfected species.

**Table A.4** Comparison of *Wolbachia* infected individuals sampled from equatorial and tropical regions for **(A)** *B. bryoniae*, **(B)** *B. frauenfeldi*, **(C)** *B. neohumeralis,* **(D)** *B. strigifinis* and **(E)** *B. tryoni* over five collection years. Fisher's exact test shows overall no significant (ns) temporal effect at α = 0.05, except for *B. tryoni,* but this did not represent an overall increase or decrease in prevalence.

| | **A** | **B** | **C** | **D** | **E** |
|---|---|---|---|---|---|
| Collection year | *B. bryoniae* | *B. frauenfeldi* | *B. neohumeralis* | *B. strigifinis* | *B. tryoni* |
| 1997 | 0/7 | | | | |
| 1998 | 1/8 | 0/3 | 2/14 | 3/12 | 0/23 |
| 2001 | 1/2 | 2/6 | 1/4 | 0/3 | 2/5 |
| 2012 | | 0/5 | 10/64 | 0/4 | |
| 2013 | 2/28 | 3/20 | | 2/18 | 2/65 |
| p value | p = 0.222 | p=0.612 | p = 0.848 | p = 0.610 | p = 0.019* |

* significant at p < 0.05

**Table A.5** Distance matrix of *COI* sequences covering 551bp positions of *Bactrocera tryoni* and *Bactrocera neohumeralis* individuals, grouped by *Wolbachia*-infected (grey shading) and uninfected (no shading). Below the diagonal are the pairwise number of differences and above the diagonal are p-distance values. P-distances given for within- and between- group means below the main table.

| Species and individual ID. No. | *B. neohumeralis_106* | *B. neohumeralis_142* | *B. neohumeralis_143* | *B. neohumeralis_23* | *B. neohumeralis_24* | *B. neohumeralis_242* | *B. neohumeralis_26* | *B. neohumeralis_27* | *B. tryoni_121* | *B. tryoni_152* | *B. tryoni_167* | *B. tryoni_182* | *B. neohumeralis_109* | *B. neohumeralis_160* | *B. neohumeralis_221* | *B. neohumeralis_238* | *B. neohumeralis_240* | *B. neohumeralis_243* | *B. neohumeralis_244* | *B. neohumeralis_248* | *B. neohumeralis_35* | *B. tryoni_275* | *B. tryoni_276* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *B. neohumeralis_106* | | 0.016 | 0.016 | 0.018 | 0.027 | 0.016 | 0.022 | 0.016 | 0.011 | 0.020 | 0.015 | 0.016 | 0.016 | 0.025 | 0.020 | 0.018 | 0.018 | 0.022 | 0.025 | 0.016 | 0.025 | 0.009 | 0.016 |
| *B. neohumeralis_142* | 9 | | 0.004 | 0.005 | 0.018 | 0.004 | 0.013 | 0.004 | 0.009 | 0.011 | 0.016 | 0.018 | 0.004 | 0.016 | 0.011 | 0.009 | 0.005 | 0.016 | 0.016 | 0.007 | 0.016 | 0.011 | 0.007 |
| *B. neohumeralis_143* | 9 | 2 | | 0.005 | 0.018 | 0.004 | 0.013 | 0.004 | 0.009 | 0.011 | 0.016 | 0.018 | 0.004 | 0.016 | 0.011 | 0.009 | 0.005 | 0.016 | 0.016 | 0.007 | 0.016 | 0.011 | 0.007 |
| *B. neohumeralis_23* | 10 | 3 | 3 | | 0.016 | 0.005 | 0.011 | 0.005 | 0.011 | 0.013 | 0.018 | 0.020 | 0.005 | 0.015 | 0.009 | 0.007 | 0.004 | 0.016 | 0.015 | 0.009 | 0.015 | 0.013 | 0.009 |
| *B. neohumeralis_24* | 15 | 10 | 10 | 9 | | 0.018 | 0.005 | 0.015 | 0.020 | 0.018 | 0.024 | 0.025 | 0.018 | 0.005 | 0.007 | 0.009 | 0.016 | 0.011 | 0.009 | 0.018 | 0.002 | 0.022 | 0.022 |
| *B. neohumeralis_242* | 9 | 2 | 2 | 3 | 10 | | 0.013 | 0.004 | 0.009 | 0.011 | 0.016 | 0.018 | 0.004 | 0.016 | 0.011 | 0.009 | 0.005 | 0.016 | 0.016 | 0.007 | 0.016 | 0.007 | 0.007 |
| *B. neohumeralis_26* | 12 | 7 | 7 | 6 | 3 | 7 | | 0.009 | 0.015 | 0.013 | 0.018 | 0.020 | 0.013 | 0.004 | 0.002 | 0.004 | 0.011 | 0.005 | 0.004 | 0.013 | 0.007 | 0.016 | 0.016 |
| *B. neohumeralis_27* | 9 | 2 | 2 | 3 | 8 | 2 | 5 | | 0.009 | 0.007 | 0.013 | 0.015 | 0.004 | 0.013 | 0.007 | 0.005 | 0.005 | 0.013 | 0.013 | 0.007 | 0.013 | 0.011 | 0.007 |
| *B. tryoni_121* | 6 | 5 | 5 | 6 | 11 | 5 | 8 | 5 | | 0.013 | 0.018 | 0.020 | 0.009 | 0.018 | 0.013 | 0.011 | 0.011 | 0.018 | 0.018 | 0.009 | 0.018 | 0.013 | 0.013 |
| *B. tryoni_152* | 11 | 6 | 6 | 7 | 10 | 6 | 7 | 4 | 7 | | 0.016 | 0.018 | 0.011 | 0.016 | 0.011 | 0.009 | 0.013 | 0.016 | 0.016 | 0.004 | 0.016 | 0.015 | 0.015 |
| *B. tryoni_167* | 8 | 9 | 9 | 10 | 13 | 9 | 10 | 7 | 10 | 9 | | 0.002 | 0.016 | 0.022 | 0.016 | 0.015 | 0.018 | 0.022 | 0.022 | 0.016 | 0.022 | 0.009 | 0.020 |
| *B. tryoni_182* | 9 | 10 | 10 | 11 | 14 | 10 | 11 | 8 | 11 | 10 | 1 | | 0.018 | 0.024 | 0.018 | 0.016 | 0.020 | 0.024 | 0.024 | 0.018 | 0.024 | 0.011 | 0.022 |
| *B. neohumeralis_109* | 9 | 2 | 2 | 3 | 10 | 2 | 7 | 2 | 5 | 6 | 9 | 10 | | 0.013 | 0.011 | 0.009 | 0.005 | 0.016 | 0.016 | 0.007 | 0.016 | 0.011 | 0.007 |
| *B. neohumeralis_160* | 14 | 9 | 9 | 8 | 3 | 9 | 2 | 7 | 10 | 9 | 12 | 13 | 7 | | 0.005 | 0.007 | 0.015 | 0.009 | 0.007 | 0.016 | 0.007 | 0.020 | 0.020 |
| *B. neohumeralis_221* | 11 | 6 | 6 | 5 | 4 | 6 | 1 | 4 | 7 | 6 | 9 | 10 | 6 | 3 | | 0.002 | 0.009 | 0.007 | 0.005 | 0.011 | 0.009 | 0.015 | 0.015 |
| *B. neohumeralis_238* | 10 | 5 | 5 | 4 | 5 | 5 | 2 | 3 | 6 | 5 | 8 | 9 | 5 | 4 | 1 | | 0.007 | 0.009 | 0.007 | 0.009 | 0.011 | 0.013 | 0.013 |
| *B. neohumeralis_240* | 10 | 3 | 3 | 2 | 9 | 3 | 6 | 3 | 6 | 7 | 10 | 11 | 3 | 8 | 5 | 4 | | 0.016 | 0.015 | 0.009 | 0.015 | 0.013 | 0.005 |
| *B. neohumeralis_243* | 12 | 9 | 9 | 9 | 6 | 9 | 3 | 7 | 10 | 9 | 12 | 13 | 9 | 5 | 4 | 5 | 9 | | 0.009 | 0.016 | 0.013 | 0.020 | 0.020 |
| *B. neohumeralis_244* | 14 | 9 | 9 | 8 | 5 | 9 | 2 | 7 | 10 | 9 | 12 | 13 | 9 | 4 | 3 | 4 | 8 | 5 | | 0.016 | 0.011 | 0.020 | 0.020 |
| *B. neohumeralis_248* | 9 | 4 | 4 | 5 | 10 | 4 | 7 | 4 | 5 | 2 | 9 | 10 | 4 | 9 | 6 | 5 | 5 | 9 | 9 | | 0.016 | 0.011 | 0.011 |
| *B. neohumeralis_35* | 14 | 9 | 9 | 8 | 1 | 9 | 4 | 7 | 10 | 9 | 12 | 13 | 9 | 4 | 5 | 6 | 8 | 7 | 6 | 9 | | 0.020 | 0.020 |
| *B. tryoni_275* | 5 | 6 | 6 | 7 | 12 | 4 | 9 | 6 | 7 | 8 | 5 | 6 | 6 | 11 | 8 | 7 | 7 | 11 | 11 | 6 | 11 | | 0.015 |
| *B. tryoni_276* | 9 | 4 | 4 | 5 | 12 | 4 | 9 | 4 | 7 | 8 | 11 | 12 | 4 | 11 | 8 | 7 | 3 | 11 | 11 | 6 | 11 | 8 | |

| Group means | within group | between groups |
|---|---|---|
| *Wolbachia* negative | 0.0136 | |
| *Wolbachia* positive | 0.0122 | 0.0131 |

# Appendix B

**Table B.1** Collection data for 104 field-caught, male fruit flies from this study, including collection location and year, specimen specific ID no. and result for *wsp,* 16S rDNA and MLST screening.

| | | | | | | | | MLST | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Location | | | | | | | standard (F1-R1) including reamplification | | | | | nested (F3-R3/F1-R1) | | | Bspecific | | | | qPCR | | |
| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | ID No. | *wsp* | 16S | *ftsZ* | *coxA* | *fbpA* | *gatB* | *hcpA* | *fbpA* | *gatB* | *hcpA* | *coxA* | *gatB* | *hcpA* | *fbpA* | *wsp* | *fbpA* | *coxA* |
| *B. bryoniae* | Weipa | 12.63 | 141.92 | 1997 | 272 | n | n | | | | | | | | | | | | | | n | |
| *B. bryoniae* | Weipa | 12.63 | 141.92 | 1997 | 273 | n | n | | | | | | | | | | | | | | n | |
| *B. bryoniae* | Weipa | 12.63 | 141.92 | 1997 | 274 | n | n | | | | | | | | | | | | | | n | |
| *B. bryoniae* | Lockhart River | 12.79 | 143.34 | 2013 | 536 | y | y | | y | | y | | y | | y | | | | | | | |
| *B. bryoniae* | Kewarra Beach | 16.78 | 145.68 | 2013 | 545 | y | y | y | | | | | | | | | | | | | | |
| *B. bryoniae* | Edge Hill | 16.90 | 145.74 | 2001 | 157 | y | y | y | y | | | y | y | y | y | | | | | y | y | y |
| *B. bryoniae* | Gordonvale | 17.07 | 145.77 | 1998 | 146 | y | n | n | n | n | n | n | n | n | y | | | n | n | y | y | y |
| *B. cacuminata* | Richmond | 33.60 | 150.75 | 2011 | 186 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. cacuminata* | Richmond | 33.60 | 150.75 | 2011 | 187 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. cacuminata* | Richmond | 33.60 | 150.75 | 2011 | 188 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. decurtans* | Seisia | 10.85 | 142.35 | 1998 | 85 | y | y | y | y | y | y | y | | y | y | | | | | y | y | y |
| *B. decurtans* | Seisia | 10.85 | 142.35 | 1998 | 264 | n | n | | | | | | | | | | | | | | n | |
| *B. decurtans* | Seisia | 10.85 | 142.35 | 1998 | 265 | n | n | | | | | | | | | | | | | | n | |
| *B. decurtans* | Seisia | 10.85 | 142.35 | 1998 | 266 | n | n | | | n | | n | n | | n | | | | | | | |
| *B. frauenfeldi* | Seisia | 10.85 | 142.35 | 2012 | 191 | n | n | n | n | n | n | n | n | n | n | | | | | | | |
| *B. frauenfeldi* | Kewarra Beach | 16.78 | 145.68 | 2013 | 485 | y | y | y | y | | y | | y | | y | | | | | | | |
| *B. frauenfeldi* | Freshwater | 16.88 | 145.72 | 2013 | 490 | y | y | y | y | | | | y | | | | | | | | | |
| *B. frauenfeldi* | Freshwater | 16.88 | 145.72 | 2013 | 491 | n | n | | | | n | n | | | | | | | | | n | n |
| *B. frauenfeldi* | Freshwater | 16.88 | 145.72 | 2013 | 492 | y | y | y | y | | y | | y | | y | | | | | | | |
| *B. frauenfeldi* | Edge Hill | 16.90 | 145.74 | 2001 | 158 | y | n | n | n | n | n | n | y | n | n | | | n | n | | y | y |
| *B. frauenfeldi* | Cairns | 16.92 | 145.78 | 2001 | 136 | y | y | y | y | n | n | y | y | y | y | | | | | | | |
| *B. neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | 204 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | 214 | n | n | n | n | n | | | n | | | | | | | | | |
| *B. neohumeralis* | Lockerbie | 10.47 | 142.27 | 2012 | 215 | n | n | n | n | n | | | n | | | | | | | | | |

| Species | Location | | | Collection Year | ID No. | *wsp* | 16S | MLST | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | standard (F1-R1) including reamplification | | | | | nested (F3-R3/F1-R1) | | | Bspecific | | | | qPCR | | |
| | Locale | Latitude (°S) | Longitude (°E) | | | | | *ftsZ* | *coxA* | *fbpA* | *gatB* | *hcpA* | *fbpA* | *gatB* | *hcpA* | *coxA* | *gatB* | *hcpA* | *fbpA* | *wsp* | *fbpA* | *coxA* |
| *B. neohumeralis* | Seisia | 10.85 | 142.35 | 2012 | 193 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. neohumeralis* | Seisia | 10.85 | 142.35 | 2012 | 194 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. neohumeralis* | Kewarra Beach | 16.79 | 145.68 | 2012 | 242 | n | n | | n | | | | | | | | | | | | n | n |
| *B. neohumeralis* | Kewarra Beach | 16.79 | 145.68 | 2012 | 243 | y | y | y | y | | | n | y | y | y | | | | | y | y | y |
| *B. neohumeralis* | Kewarra Beach | 16.79 | 145.68 | 2012 | 244 | y | y | y | y | | | n | y | y | n | | | | | | | |
| *B. neohumeralis* | Freshwater | 16.88 | 145.71 | 2012 | 221 | y | y | y | y | y | | | y | y | y | | | | | | | |
| *B. neohumeralis* | Freshwater | 16.88 | 145.71 | 2012 | 248 | y | y | y | y | | | y | y | y | | | | | | y | y | y |
| *B. neohumeralis* | Cairns | 16.92 | 145.78 | 2001 | 160 | y | n | n | n | n | n | n | | n | n | | | n | n | y | y | y |
| *B. neohumeralis* | Bayview Heights | 16.96 | 145.72 | 1998 | 35 | y | y | y | y | n | n | n | y | y | y | | | | | y | | |
| *B. neohumeralis* | Bayview Heights | 16.96 | 145.72 | 1998 | 109 | y | y | y | y | n | n | n | y | y | y | | | | | y | y | y |
| *B. neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | 106 | n | n | | n | | n | n | | n | n | | | n | n | | | |
| *B. neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | 142 | n | n | n | n | n | n | n | | n | n | | | n | n | n | n | n |
| *B. neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | 143 | n | n | n | n | n | n | n | | n | n | | | n | n | n | n | n |
| *B. neohumeralis* | Gordonvale | 17.07 | 145.77 | 1998 | 145 | n | n | n | n | n | n | n | | n | n | | | n | n | n | n | n |
| *B. neohumeralis* | Mourilyan Harbour | 17.35 | 146.40 | 2012 | 238 | y | y | y | y | y | | | y | y | y | | | | | y | | |
| *B. neohumeralis* | Mourilyan Harbour | 17.36 | 146.80 | 2012 | 240 | y | y | y | y | y | | | y | y | y | | | | | y | y | y |
| *B. neohumeralis* | Townsville | 19.36 | 146.84 | 2012 | 381 | n | n | | n | | | | n | | n | | | | | | | |
| *B. neohumeralis* | Mackay | 21.08 | 149.19 | 2012 | 342 | y | y | y | y | | y | | n | | | | | | | | | |
| *B. neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | 345 | y | y | y | y | | | | n | | y | | | | | | | |
| *B. neohumeralis* | Mackay | 21.11 | 149.10 | 2012 | 346 | y | y | y | y | | y | | y | | y | | | | | | | |
| *B. neohumeralis* | Mackay | 21.14 | 149.18 | 2012 | 230 | n | n | n | n | n | | | n | | | | | | | | | |
| *B. neohumeralis* | Gladstone | 23.88 | 151.24 | 2012 | 355 | y | y | | y | | y | | y | | y | | | | | | | |
| *B. neohumeralis* | Sunnybank | 27.58 | 153.06 | 1998 | 26 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. neohumeralis* | Sunnybank | 27.58 | 153.06 | 1998 | 27 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | 24 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. neohumeralis* | Lismore | 28.82 | 153.28 | 1998 | 25 | n | n | n | n | n | n | n | n | n | n | | | | | | | |
| *B. neohumeralis* | North Sapphire | 30.23 | 153.15 | 1998 | 22 | n | n | | | | | | | n | | | | | | | | |

| Species | Location | | | Collection Year | ID No. | *wsp* | 16S | MLST | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | standard (F1-R1) including reamplification | | | | | nested (F3-R3/F1-R1) | | | Bspecific | | | | qPCR | | |
| | Locale | Latitude (°S) | Longitude (°E) | | | | | *ftsZ* | *coxA* | *fbpA* | *gatB* | *hcpA* | *fbpA* | *gatB* | *hcpA* | *coxA* | *gatB* | *hcpA* | *fbpA* | *wsp* | *fbpA* | *coxA* |
| *B. neohumeralis* | North Sapphire | 30.23 | 153.15 | 1998 | 23 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. peninsularis* | Cape York | 10.58 | 142.22 | 1997 | 71 | n | y | n | | n | n | n | n | n | n | | | n | y | | | |
| *B. peninsularis* | Cape York | 10.58 | 142.22 | 1997 | 72 | n | y | n | | n | n | n | n | n | n | | | n | y | | | |
| *B. peninsularis* | Cape York | 10.58 | 142.22 | 1997 | 73 | n | y | n | | n | n | n | n | n | n | | | n | y | | | |
| *B. peninsularis* | Seisia | 10.85 | 142.35 | 2012 | 196 | n | y | n | n | | n | | | n | | | | n | y | | | |
| *B. peninsularis* | Seisia | 10.85 | 142.35 | 2012 | 197 | n | y | n | | | n | | | n | | | | n | y | | | |
| *B. peninsularis* | Edge Hill | 16.90 | 145.74 | 2001 | 153 | n | y | n | n | | n | n | | n | n | n | n | | | | | |
| *B. perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | 74 | y | y | n | y | y | y | y | | y | y | | | | | | | |
| *B. perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | 75 | y | y | n | y | y | y | y | | y | y | | | | | | | |
| *B. perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | 76 | y | y | n | y | y | y | y | | y | y | | | | | | | |
| *B. perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | 261 | y | y | n | y | | | y | y | y | | | | | | | | |
| *B. perkinsi* | Roma Flats | 10.45 | 142.31 | 1997 | 263 | y | y | n | y | | | y | y | y | | | | | | | | |
| *B. quadrata* | Bayview Heights | 16.96 | 145.72 | 1998 | 149 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | 81 | y | y | n | n | n | n | n | y | n | y | | | | | y | y | y |
| *B. strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | 269 | y | y | y | y | | | n | y | y | y | | | | | y | y | y |
| *B. strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | 270 | n | n | | | | | | | | | | | | | | n | |
| *B. strigifinis* | Bamaga | 10.89 | 142.39 | 1998 | 271 | n | n | | | | | | | | | | | | | | n | |
| *B. strigifinis* | Lockhart River | 12.60 | 143.41 | 2013 | 517 | n | n | n | | | n | | n | | | | | | | | | |
| *B. strigifinis* | Kewarra Beach | 16.78 | 145.68 | 2013 | 503 | y | y | y | y | | y | | y | | y | | | | | | | |
| *B. strigifinis* | Kewarra Beach | 16.78 | 145.68 | 2013 | 504 | y | y | y | y | | n | | y | | | | | | | | | |
| *B. strigifinis* | Cairns | 16.92 | 145.78 | 2001 | 159 | n | n | n | n | n | n | n | n | n | n | | | | | | | |
| *B. strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | 138 | y | n | n | n | n | n | n | n | n | y | | | n | n | y | n | n |
| *B. strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | 140 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. strigifinis* | Gordonvale | 17.07 | 145.77 | 1998 | 141 | n | n | n | n | n | n | n | n | n | n | | | n | n | n | n | n |
| *B. tryoni* | Edge Hill | 16.90 | 145.74 | 2001 | 152 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | n |
| *B. tryoni* | Cairns | 16.92 | 145.78 | 2001 | 275 | y | y | y | y | | | | y | y | y | | | | | y | y | y |
| *B. tryoni* | Cairns | 16.92 | 145.78 | 2001 | 276 | y | y | y | y | | | | y | y | y | | | | | y | y | y |

| | Location | | | | | | | MLST | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | standard (F1-R1) including reamplification | | | | | nested (F3-R3/F1-R1) | | | Bspecific | | | | qPCR | | |
| Species | Locale | Latitude (°S) | Longitude (°E) | Collection Year | ID No. | *wsp* | 16S | *ftsZ* | *coxA* | *fbpA* | *gatB* | *hcpA* | *fbpA* | *gatB* | *hcpA* | *coxA* | *gatB* | *hcpA* | *fbpA* | *wsp* | *fbpA* | *coxA* |
| *B. tryoni* | Mackay | 21.08 | 149.19 | 2013 | 443 | y | y | y | y | | | | y | | | | | | | | | |
| *B. tryoni* | Mackay | 21.11 | 149.10 | 2013 | 439 | y | y | y | y | | | | | | | | | | | | | |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 164 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 165 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 167 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | n |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 169 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 170 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | n |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 173 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | n |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 174 | n | n | | | | n | n | | | | | | n | n | n | n | |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 180 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | n |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 181 | n | n | n | n | n | n | n | | n | n | n | n | n | n | n | n | |
| *B. tryoni* | Alice Springs | 23.70 | 133.88 | 1998 | 182 | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n |
| *B. tryoni* | Brisbane | 27.66 | 152.98 | 2013 | 400 | n | n | | n | | | | n | | n | | | | | | | |
| *B. tryoni* | Brisbane | 27.66 | 152.98 | 2013 | 403 | n | n | | n | | | | n | | n | | | | | | | |
| *B. visenda* | Pajinka | 10.42 | 142.32 | 2012 | 211 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. visenda* | Injinoo | 10.54 | 142.29 | 2012 | 208 | n | n | | | | n | n | | n | n | | | | | | | |
| *B. visenda* | New Mapoon | 10.87 | 142.39 | 2012 | 209 | n | n | n | n | n | n | n | n | n | n | | | | | | | |
| *B. visenda* | New Mapoon | 10.87 | 142.39 | 2012 | 210 | n | n | | | | n | n | | n | n | | | | | | | |
| *D. aequalis* | Narara | 33.40 | 151.35 | 1998 | 52 | n | n | | n | | n | n | | n | n | | | | | | | |
| *D. axanus* | Weipa | 12.63 | 141.92 | 1998 | 88 | y | y | y | y | n | n | y | y | y | y | | | | | y | y | y |
| *D. axanus* | Weipa | 12.63 | 141.92 | 1998 | 267 | n | n | | | | n | n | | | | | | | | | n | |
| *D. axanus* | Weipa | 12.63 | 141.92 | 2001 | 268 | n | n | | | | n | n | | | | | | | | | n | |
| *D. newmani* | Alice Springs | 23.70 | 133.88 | 1998 | 168 | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | |
| *D. newmani* | Alice Springs | 23.70 | 133.88 | 1998 | 176 | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | |
| *D. newmani* | Alice Springs | 23.70 | 133.88 | 1998 | 177 | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | |
| *D. newmani* | Alice Springs | 23.70 | 133.88 | 1998 | 185 | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | n | |

**Table B.2** *Wolbachia*, mitochondrial and qPCR primers.

| Locus | Primer name | Sequence (5'-3') | Reference |
|---|---|---|---|
| *Wolbachia* | | | |
| *wsp* | Wsp-F | TGGTCCAATAAGTGATGAAGAAACTAGCTA | Jeyaprakash and Hoy, 2000 |
| | Wsp-R | AAAAATTAAACGCTACTCCAGCTTCTGCAC | Jeyaprakash and Hoy, 2000 |
| 16S | wspecF | CATACCTATTCGAAGGGATAG | Werren and Windsor, 2000 |
| | wspecR | AGCTTCGAGTGAAACCAATTC | Werren and Windsor, 2000 |
| *gatB* | gatB_F3 | ATTCAYYTAGARCAAGATGCAGG | Baldo et al. 2006 |
| | gatB_R3 | AAGAGCTCKGAYAAAGCATYBGC | Baldo et al. 2006 |
| | gatB_F1 | GAKTTAAAYCGYGCAGGBGTT | Baldo et al. 2006 |
| | gatB_R1 | TGGYAAYTCRGGYAAAGATGA | Baldo et al. 2006 |
| | gatB_BspecF1 | TAAGAATCGCAAGAATTCAC | Baldo et al. 2006 |
| *coxA* | coxA_F1 | TTGGRGCRATYAACTTTATAG | Baldo et al. 2006 |
| | coxA_R1 | CTAAAGACTTTKACRCCAGT | Baldo et al. 2006 |
| | coxA_BspecF1 | ATACCCACCTYTRTCGCAAA | Baldo et al. 2006 |
| *hcpA* | hcpA_F3 | ATTAGAGAAATARCAGTTGCTGC | Baldo et al. 2006 |
| | hcpA_R3 | CATGAAAGACGAGCAARYTCTGG | Baldo et al. 2006 |
| | hcpA_F1 | GAAATARCAGTTGCTGCAAA | Baldo et al. 2006 |
| | hcpA_R1 | GAAAGTYRAGCAAGYTCTG | Baldo et al. 2006 |
| | hcpA_BspecR1 | TTCTTTGTCGCTMACTTYAATCAKG | Baldo et al. 2006 |
| *ftsZ* | ftsZ_F1 | ATYATGGARCATATAAARGATAG | Baldo et al. 2006 |
| | ftsZ_R1 | TCRAGYAATGGATTRGATAT | Baldo et al. 2006 |
| *fbpA* | fbpA_F3 | GTTAACCCTGATGCYYAYGAYCC | Baldo et al. 2006 |
| | fbpA_R3 | TCTACTTCCTTYGAYTCDCCRCC | Baldo et al. 2006 |
| | fbpA_F1 | GCTGCTCCRCTTGGYWTGAT | Baldo et al. 2006 |
| | fbpA_R1 | CCRCCAGARAAAAYYACTATTC | Baldo et al. 2006 |
| | fbpA_BspecF1 | GTTAACCCTGATGCTTACGAT | Baldo et al. 2006 |
| Mitochondrial | | | |
| *COI* | Dick (C1-J-2441) | CCTACAGGAATTAAAATTTTTAGATGATTA | Simon et al. 1994 |
| | Pat (TL2-N-3014) | TCCATTGCACTAATCTGCCATATTA | Simon et al. 1994 |
| | LCO1490 | GGTCAACAAATCATAAAGATATTGG | Folmer et al. 1994 |
| | HCO2198 | TAAACTTCAGGGTGACCAAAAAATCA | Folmer et al. 1994 |
| qPCR | | | |
| *wsp* | wsp11_F | CTC CGG AAG TCA AAC TTT ATG CTG G | this study |
| | wsp11_R | TGT CTT TGC CTG CAG CAG CAT CTT | this study |
| | wsp661_F | GCT ACG ACG TAA CTC CAG AAA TCA AA | this study |
| | wsp661_R | TTT ACC TGC CGC ACC AGC AGT TT | this study |
| *coxA* | cox15_F | GCA TGT CAT TAA CTA AGA TGC CAC TGT | this study |
| | cox15_R | AAT AAC ACA GGA TCG CCA CCA CCG | this study |
| | cox84_F | ATG TCG TTG ACT AAG ATG CCA CTA T | this study |
| | cox84_R | AAT AAC ACA GGG TCA CCA CCA CCT G | this study |
| *fbpA* | fbp62_F | GTT TGG GAT GCT TAG CTG TCG GAT | this study |
| | fbp62_R | TAT GAT TCC ACG GGC TTC CTC CAT | this study |
| | fbp17_F | GGC TGC TTG GCT GTT GGA TTT ACT | this study |
| | fbp17_R | TCA GCT ACG ATT TCA CGG GCT TCT | this study |
| *scarlet* | scarBt_F | GCC ACA TTC TTC GCC TTC AGC ATA | this study |
| | scarBt_R | TAA TCG ACG GGC ACC AAA TAA GCC | this study |
| | scarDax_F | GCC ACA TTC TTC GCG TTT AGC ATA | this study |
| | scarDax_R | TAA TCG ACA GGC ACC AAG TAA GCC | this study |

**Table B.3** Number of *Wolbachia* isolates in the MLST database (645 isolates of 362 STs; accessed on 12 December 2013) that share MLST alleles with ST-17 and ST-370 isolated from *B. frauenfeldi, B. cacuminata* and *F. arisanus*. The detection frequency of alleles and MLST allele combinations of ST-17 and ST-370 in three out of 26 host species in our study is significantly different from what could be expected by binomial sampling of the MLST database.

| | *gatB* | *coxA* | *hcpA* | *ftsZ* | *fbpA* | MLST allele combination |
|---|---|---|---|---|---|---|
| ST-17 | 6 | 3 | 3 | 64 | 4 | |
| p | ** | *** | *** | ns | *** | *** |
| ST-370 | 1 | 1 | 3 | 4 | 1 | |
| p | *** | *** | *** | *** | *** | *** |

***p < 0.001; **p < 0.01; *p < 0.05; ns = not significant

**Table B.4** *Wolbachia* strain density values. Mean quantification cycle values (Cq) and individual data points normalised to the *scarlet* reference gene. Letters a-e for wsp11 and f-j for wsp661 refer to significantly different means, also see Figure 3.1 and Figure 3.2.

| Individual | Allele | Cq | normalised mean ± s.d (CV) | | | | Tukey's Test* | |
|---|---|---|---|---|---|---|---|---|
| | | | 2^-(CqTarget-CqRef) | | | | wsp11 | wsp661 |
| Bb157 | wsp11 | 24.61 | 0.00850 | ± | 0.00057 | (6.73) | e | |
| | wsp661 | 26.71 | 0.00198 | ± | 0.00002 | (0.98) | | j |
| | fbp17 | 24.91 | 0.00694 | ± | 0.00111 | (16.0) | | |
| | fbp62 | 27.58 | 0.00132 | ± | 0.00011 | (8.32) | | |
| | cox15 | 23.36 | 0.03675 | ± | 0.00117 | (3.17) | | |
| | cox84 | 25.77 | 0.00702 | ± | 0.00147 | (21.0) | | |
| Bd85 | wsp11 | 18.25 | 1.82935 | ± | 0.16908 | (9.24) | b | |
| | wsp661 | 19.17 | 0.96944 | ± | 0.01896 | (1.96) | | g |
| | fbp17 | 18.65 | 1.39478 | ± | 0.20921 | (15.0) | | |
| | fbp62 | 19.92 | 0.58281 | ± | 0.10765 | (18.5) | | |
| | cox15 | 18.24 | 1.86449 | ± | 0.33271 | (17.8) | | |
| | cox84 | 19.06 | 1.05545 | ± | 0.16281 | (15.4) | | |
| Bn35 | wsp11 | 29.43 | 0.00081 | ± | 0.00002 | (2.45) | e | |
| | wsp661 | 28.73 | 0.00134 | ± | 0.00027 | (19.9) | | j |
| Bn109 | wsp11 | 27.63 | 0.00450 | ± | 0.00007 | (1.47) | e | |
| | wsp661 | 28.28 | 0.00285 | ± | 0.00008 | (2.81) | | j |
| | fbp17 | 28.92 | 0.00147 | ± | 0.00022 | (14.8) | | |
| | fbp62 | 29.85 | 0.00097 | ± | 0.00016 | (16.8) | | |
| Bn221 | wsp11 | 15.85 | 6.24877 | ± | 0.14002 | (2.24) | a | |
| | wsp661 | 17.36 | 2.21452 | ± | 0.38358 | (17.3) | | f |
| | fbp17 | 16.08 | 5.32741 | ± | 0.07663 | (1.44) | | |
| | fbp62 | 17.77 | 1.64477 | ± | 0.08361 | (5.08) | | |
| | cox15 | 15.18 | 10.0604 | ± | 1.99242 | (19.8) | | |
| | cox84 | 16.31 | 4.53219 | ± | 0.09424 | (2.08) | | |
| Bn238 | wsp11 | 19.41 | 0.31427 | ± | 0.00376 | (1.20) | d | |
| | wsp661 | 19.15 | 0.38071 | ± | 0.06132 | (16.1) | | i |
| Bn240 | wsp661 | 21.18 | 0.17083 | ± | 0.00622 | (3.64) | | ij |
| | fbp62 | 21.73 | 0.11688 | ± | 0.00929 | (7.94) | | |
| | cox84 | 21.02 | 0.19052 | ± | 0.01014 | (5.32) | | |
| Bn243 | wsp661 | 25.19 | 0.00516 | ± | 0.00047 | (9.04) | | j |
| | fbp62 | 25.67 | 0.00375 | ± | 0.00080 | (18.9) | | |
| | cox84 | 24.26 | 0.00993 | ± | 0.00187 | (21.4) | | |
| Bn244 | wsp11 | 27.07 | 0.00241 | ± | 0.00084 | (35.0) | e | |
| | fbp17 | 27.71 | 0.00150 | ± | 0.00008 | (5.39) | | |
| Bn248 | wsp11 | 17.03 | 0.74788 | ± | 0.02870 | (3.84) | c | |
| | wsp661 | 17.28 | 0.62927 | ± | 0.04701 | (7.47) | | h |
| | fbp17 | 17.23 | 0.64807 | ± | 0.04125 | (6.36) | | |
| | fbp62 | 17.86 | 0.42926 | ± | 0.09948 | (23.2) | | |
| | cox15 | 16.46 | 1.10945 | ± | 0.06255 | (5.63) | | |
| | cox84 | 16.76 | 0.90501 | ± | 0.16067 | (17.8) | | |

| Individual | Allele | Cq | normalised mean ± s.d (CV) | | | | Tukey's Test* | |
|---|---|---|---|---|---|---|---|---|
| | | | 2^-(CqTarget-CqRef) | | | | wsp11 | wsp661 |
| Bs81 | wsp11 | 29.44 | 0.00375 | ± | 0.00140 | (37.3) | e | |
| | wsp661 | 30.82 | 0.00140 | ± | 0.00018 | (12.7) | | j |
| Bs269 | wsp11 | 28.53 | 0.00241 | ± | 0.00044 | (18.1) | e | |
| | wsp661 | 29.7 | 0.00107 | ± | 0.00019 | (17.6) | | j |
| | fbp17 | 28.38 | 0.00269 | ± | 0.00048 | (17.7) | | |
| | fbp62 | 29.33 | 0.00138 | ± | 0.00016 | (11.8) | | |
| Bt275 | wsp661 | 23.93 | 0.00780 | ± | 0.00066 | (8.44) | | j |
| | fbp62 | 24.48 | 0.00534 | ± | 0.00077 | (14.4) | | |
| | cox84 | 23.43 | 0.01100 | ± | 0.00016 | (1.44) | | |
| Bt276 | wsp661 | 26.43 | 0.00091 | ± | 0.00002 | (2.45) | | j |
| | fbp62 | 26.69 | 0.00077 | ± | 0.00016 | (21.1) | | |
| | cox84 | 25.28 | 0.00202 | ± | 0.00006 | (2.77) | | |
| Dax88 | wsp11 | 25.47 | 0.01027 | ± | 0.00073 | (7.09) | e | |
| | wsp661 | 26.13 | 0.00659 | ± | 0.00131 | (19.8) | | j |
| | fbp17 | 25.49 | 0.01027 | ± | 0.00198 | (19.2) | | |
| | fbp62 | 26.57 | 0.00483 | ± | 0.00076 | (15.7) | | |
| | cox15 | 24.2 | 0.02436 | ± | 0.00084 | (3.47) | | |
| | cox84 | 25.45 | 0.01024 | ± | 0.00083 | (8.09) | | |

**Table B.5** ANOVA of the linkage of genes within each *Wolbachia* chromosome. Hypothetical combinations of pairwise linkage were examined for each individual; a significant difference in titre is found when the linked alleles in one strain are contrasted to the linked alleles in the other strain. No significant difference is found where non-linked alleles are compared.

| | | | | | ANOVA | |
|---|---|---|---|---|---|---|
| | Gene combination (*Wolbachia* "1") | mean | Gene combination (*Wolbachia* "2") | mean | p value | significance |
| Bb157 | wsp11/fbp17 | 0.0077 | wsp661/fbp62 | 0.0017 | 0.6334 | ns |
| | wsp11/fbp62 | 0.0056 | wsp661/fbp17 | 0.0050 | 0.9943 | ns |
| | cox15/fbp17 | 0.0218 | cox84/fbp62 | 0.0047 | 0.0458 | * |
| | cox15/fbp62 | 0.0226 | cox84/fbp17 | 0.0070 | 0.0841 | ns |
| | wsp11/cox15 | 0.0226 | wsp661/cox84 | 0.0050 | 0.0288 | * |
| | wsp11/cox84 | 0.0078 | wsp11/cox15 | 0.0228 | 0.0887 | ns |
| Bd85 | wsp11/fbp17 | 1.5619 | wsp661/fbp62 | 0.7761 | <1e-04 | *** |
| | wsp11/fbp62 | 1.1494 | wsp661/fbp17 | 1.2125 | 0.944 | ns |
| | cox15/fbp17 | 1.6296 | cox84/fbp62 | 0.8529 | <0.001 | *** |
| | cox15/fbp62 | 1.3152 | cox84/fbp17 | 1.2251 | 0.876 | ns |
| | wsp11/cox15 | 1.8510 | wsp661/cox84 | 1.0186 | <1e-04 | *** |
| | wsp11/cox84 | 1.3531 | wsp11/cox15 | 1.4809 | 0.7589 | ns |
| Bn221 | wsp11/fbp17 | 5.7881 | wsp661/fbp62 | 1.9296 | 0.0093 | ** |
| | wsp11/fbp62 | 3.9468 | wsp661/fbp17 | 3.7710 | 0.9924 | ns |
| | cox15/fbp17 | 7.6939 | cox84/fbp62 | 3.0885 | 0.0089 | ** |
| | cox15/fbp62 | 5.8526 | cox84/fbp17 | 4.9298 | 0.861 | ns |
| | wsp11/cox15 | 8.1546 | wsp661/cox84 | 3.3734 | 0.0021 | ** |
| | wsp11/cox84 | 5.3905 | wsp11/cox15 | 6.1375 | 0.896 | ns |
| Bn248 | wsp11/fbp17 | 0.6980 | wsp661/fbp62 | 0.5293 | 0.003 | ** |
| | wsp11/fbp62 | 0.5886 | wsp661/fbp17 | 0.6387 | 0.643 | ns |
| | cox15/fbp17 | 0.8019 | cox84/fbp62 | 0.5878 | 0.073 | ns |
| | cox15/fbp62 | 0.6560 | cox84/fbp17 | 0.7337 | 0.723 | ns |
| | wsp11/cox15 | 0.8684 | wsp661/cox84 | 0.7212 | 0.143 | ns |
| | wsp11/cox84 | 0.8003 | wsp11/cox15 | 0.7893 | 0.9899 | ns |
| Bs269 | wsp11/fbp17 | 0.0025 | wsp661/fbp62 | 0.0013 | 0.0002 | *** |
| | wsp11/fbp62 | 0.0019 | wsp661/fbp17 | 0.0020 | 0.774 | ns |
| Dax88 | wsp11/fbp17 | 0.0103 | wsp661/fbp62 | 0.0057 | 0.0975 | ns |
| | wsp11/fbp62 | 0.0076 | wsp661/fbp17 | 0.0084 | 0.9244 | ns |
| | cox15/fbp17 | 0.0163 | cox84/fbp62 | 0.0071 | 0.0055 | ** |
| | cox15/fbp62 | 0.0132 | cox84/fbp17 | 0.0103 | 0.644 | ns |
| | wsp11/cox15 | 0.0163 | wsp661/cox84 | 0.0082 | 0.014 | * |
| | wsp11/cox84 | 0.0103 | wsp11/cox15 | 0.0142 | 0.418 | ns |

***p < 0.001; **p < 0.01; *p < 0.05; ns = not significant

**Table B.6** PCR conditions for *Wolbachia* detection, MLST characterisation, quantification and COI barcoding.

| **PCR protocol** | **MLST** | **qPCR** | **COI** | **COI and Sequencing** | ***wsp*** | **16S rDNA** |
|---|---|---|---|---|---|---|
| Total volume | 20µL | 10µL | 20µL | 20µL | 10µL / 20µL | 10µL |
| GoTaq Reaction Buffer | 1X | 1X | 1X | 1X | 1X | 1X |
| MgCl2 | 1.5mM | 1.5mM | 2.5mM | 2.5mM | 4mM | 2mM |
| dNTP | 200µM each | 200µM each | 200µM each | 200µM each | 150µM each | 125µM each |
| Primer 1 | 1µM | 0.4µM | 0.16µM | 0.4µM | 0.8µM | 0.63µM |
| Primer 2 | 1µM | 0.4µM | 0.16µM | 0.4µM | 0.8µM | 0.63µM |
| GoTaq DNA polymerase (Promega) | 0.5u | 0.5u | 0.5u | 0.5u | 0.4u / 0.8u | 0.5u |
| BSA | | | 125ng | | | |
| DNA Template | 2µL | 1µL | 2µL | 2µL | 1µL / 2µL | 1µL |
| **Thermal Cycling** | | | | | | |
| Denaturation | 94°C for 4min | 94°C for 4min | 94°C for 1min | 94°C for 5min | 94°C for 2min | 94°C for 5min |
| Cycling | 94°C for 30s, [primer-specific temperature] for 30s, 72°C for 1min (x35 cycles) | 94°C for 30s, 60°C for 30s, 72°C for 1min (x35 cycles) | 94°C for 30s, 45°C for 1min30s, 72°C for 1min (x5 cycles)<br>94°C for 30s, 51°C for 1min30s, 72°C for 1min (x35 cycles) | 94°C for 30s, 55°C for 30s, 72°C for 1min (x35 cycles) | 94°C for 10s, 65°C for 30s, 68°C for 1min (x10 cycles)<br>94°C for 10s, 65°C for 30s, 68°C for 1min (x25 cycles, including an incremental increase of 20s at 68°C) | 94°C for 30s, 54°C for 30s, 72°C for 1min (x35 cycles) |
| Final Extension | 72°C for 10min | 72°C for 10min | 72°C for 5min | 72°C for 10min | | 72°C for 10min |
| **Primer pairs** (annealing temp) | (Standard) | | | | | |
| | gatB_F1 / gatB_R1 (54°C) | wsp11_F / wsp11_R | LCO1490 / HCO2198 | Dick / Pat | Wsp-F / Wsp-R | wspecF / wspecR |
| | coxA_F1 / coxA_R1 (55°C) | wsp661_F / wsp661_R | | T7 Promoter / SP6 | | |
| | hcpA_F1 / hcpA_R1 (53°C) | cox15_F / cox15_R | | | | |
| | ftsZ_F1 / ftsZ_R1 (52°C) | cox84_F / cox84_R | | | | |
| | fbpA_F1 / fbpA_R1 (58°C) | fbp62_F / fbp62_R | | | | |
| | (Nested) | fbp17_F / fbp17_R | | | | |
| | gatB_F3 / gatB_R3 (55°C) | scarBt_F / scarBt_R | | | | |
| | hcpA_F3 / hcpA_R3 (55°C) | scarDax_F / scarDax_R | | | | |
| | fbpA_F3 / fbpA_R3 (55°C) | | | | | |
| | (B-group specific) | | | | | |
| | gatB_BspecF1 / gatB_R1 (62°C) | | | | | |
| | coxA_BspecF1 / coxA_R1 (54°C) | | | | | |
| | hcpA_F1 / hcpA_BspecR1 (55°C) | | | | | |
| | fbpA_BspecF1 / fbpA_R1 (58°C) | | | | | |

**Table B.7** Primer efficiency (E) values calculated from standard curves for host nuclear gene *scarlet*, and *Wolbachia* allele-specific primers.

| **Primer set** | **Template** | **Product size** | **R** | **$R^2$** | **m** | **b** | **Efficiency** | **Threshold** |
|---|---|---|---|---|---|---|---|---|
| **Reference gene** | **Genomic** | | | | | | | |
| scarBt_F - scarBt_R | *B. neohumeralis* | 137bp | 0.999 | 0.997 | -3.34 | -1.05 | 0.99 | 0.0555 |
| scarBt_F - scarBt_R | *B. strigifinis* | 137bp | 0.999 | 0.998 | -3.46 | 4.46 | 0.95 | 0.0555 |
| scarBt_F - scarBt_R | *B. bryoniae* | 137bp | 1.000 | 0.999 | -3.33 | 2.28 | 1 | 0.0555 |
| scarBt_F - scarBt_R | *B. decurtans* | 137bp | 1.000 | 0.999 | -3.39 | 2.39 | 0.97 | 0.0555 |
| scarDax_F - scarDax_R | *D. axanus* | 137bp | 0.999 | 0.998 | -3.29 | -0.54 | 1.01 | 0.0555 |
| **Target genes** | **Amplicon** | | | | | | | |
| wsp11_F - wsp11_R | wsp 11 | 87bp | 0.991 | 0.998 | -3.31 | 0.94 | 1.01 | 0.0555 |
| wsp661_F - wsp661_R | wsp 661 | 102bp | 0.999 | 0.999 | -3.56 | 0.86 | 0.98 | 0.0555 |
| cox15_F - cox15_R | cox 15 | 169bp | 1.000 | 0.999 | -3.38 | -1.26 | 0.98 | 0.0555 |
| cox84_F - cox84_R | cox 84 | 167bp | 0.999 | 0.998 | -3.22 | 0.25 | 1.04 | 0.0555 |
| fbp62_F - fbp62_R | fbp 62 | 86bp | 0.999 | 0.997 | -3.26 | 1.20 | 1.03 | 0.0555 |
| fbp17_F - fbp17_R | fbp 17 | 86bp | 0.999 | 0.999 | -3.33 | 0.95 | 1 | 0.0555 |
| **Target genes** | **Genomic** | | | | | | | |
| wsp11_F - wsp11_R | *B. decurtans* | 87bp | 0.999 | 0.998 | -3.35 | 14.94 | 0.99 | 0.0555 |
| wsp661_F - wsp661_R | *B. decurtans* | 102bp | 0.999 | 0.998 | -3.36 | 15.80 | 0.98 | 0.0555 |
| cox15_F - cox15_R | *B. decurtans* | 169bp | 0.998 | 0.996 | -3.35 | 14.54 | 0.99 | 0.0555 |
| cox84_F - cox84_R | *B. decurtans* | 167bp | 0.998 | 0.997 | -3.31 | 16.07 | 1.01 | 0.0555 |
| fbp62_F - fbp62_R | *B. decurtans* | 86bp | 0.996 | 0.992 | -3.34 | 15.97 | 0.99 | 0.0555 |
| fbp17_F - fbp17_R | *B. decurtans* | 86bp | 0.987 | 0.974 | -3.39 | 14.98 | 0.97 | 0.0555 |

# Appendix C

**Table C.1** Field and laboratory fruit fly samples.

| Field specimens | Abbreviation | Location | Host plant | Co-ordinates | Collection date | Sex | Individual ID |
|---|---|---|---|---|---|---|---|
| *B. tryoni* | BtF12 | Richmond | Tangelo | 33.35°S, 150.45°E | 2012 | F | C1 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C2 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C3 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C4 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C10 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C12 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C13 |
| *B. tryoni* | BtF12 | Richmond | Meyer lemon | 33.35°S, 150.45°E | 2012 | F | C15 |
| | | | | | | | |
| *D. pornia* | DpF12 | Richmond | Valencia orange | 33.60°S, 150.75°E | 2012 | F | A1 |
| *D. pornia* | DpF12 | Richmond | Valencia orange | 33.60°S, 150.75°E | 2012 | F | A10 |
| *D. pornia* | DpF12 | Richmond | Meyer lemon | 33.60°S, 150.75°E | 2012 | F | A19 |
| *D. pornia* | DpF12 | Richmond | Valencia orange | 33.60°S, 150.75°E | 2012 | F | A2 |
| *D. pornia* | DpF12 | Richmond | Meyer lemon | 33.60°S, 150.75°E | 2012 | F | A20 |
| *D. pornia* | DpF12 | Richmond | Tangelo | 33.60°S, 150.75°E | 2012 | F | A3 |
| *D. pornia* | DpF12 | Richmond | Tangelo | 33.60°S, 150.75°E | 2012 | F | A4 |
| *D. pornia* | DpF12 | Richmond | Joppa Orange | 33.60°S, 150.75°E | 2012 | F | A6 |
| | | | | | | | |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG01 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG02 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG03 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG04 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG05 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG06 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG07 |
| *B. cacuminata* | BcF09 | Summer Hill | Wild tobacco | 33.89°S, 151.14°E | 2009 | F | BcG08 |

| Laboratory Lines | Abbreviation | Location * | Original Source | Years in laboratory | Collection date | Sex | Individual ID |
|---|---|---|---|---|---|---|---|
| *B. tryoni* GOS | BtGWS09 | Richmond, UWS | Bathurst, NSW | 2009-2012 | 2012 | F | GOS1-GOS8 |
| *B. tryoni* HAC | BtHWS09 | Richmond, UWS | Richmond, NSW | 2009-2012 | 2012 | F | HAC1-HAC8 |
| *B. tryoni* 54 | Bt54WS10 | Richmond, UWS | Isofemale line from BtGOS | 2010-2012 | 2012 | F | Bt54.1-Bt54.8 |
| *B. neohumeralis* | BnWS09 | Richmond, UWS | Cairns, Queensland | 2009-2012 | 2012 | F | Bn1-Bn8 |
| *B. jarvisi* | BjWS10 | Richmond, UWS | Cairns, Queensland | 2010-2012 | 2012 | F | Bj1-Bj8 |
| *B. cacuminata* | BcWS09 | Richmond, UWS | Richmond, NSW | 2009-2012 | 2012 | F | Bcac1-Bcac8 |
| *B. neohumeralis* 09 | BnQD09 | Cairns, QDAFF | Cairns, Queensland | 2009-2012 | 2012 | F | Bn09.1-Bn09.8 |
| *B. neohumeralis* 12 | BnQD12 | Cairns, QDAFF | Cairns, Queensland | 2012 | 2012 | F | Bn12.1-Bn12.8 |
| *C. capitata* | CcWA99 | Perth, DAFWA | Vienna7mix99 | 1999-2011 | 2011 | F | MF1-MF8 |

*UWS (University of Western Sydney), QDAFF (Queensland Department of Agriculture, Fisheries and Forestry), DAFWA (Department of Agriculture and Food, Western Australia)

**Table C.2** Concentration and quality of pooled sample DNA for 454 pyrosequencing.

| | Nanodrop | | | | | Qubit |
|---|---|---|---|---|---|---|
| **Sample ID** | **DNA conc. (ng/μL)** | **A260** | **A280** | **260/280** | **260/230** | **DNA conc. (ng/μL)** |
| BtGWS09 | 8.5 | 0.17 | 0.071 | 2.39 | 1.15 | 23.6 |
| BtHWS09 | 22.8 | 0.456 | 0.219 | 1.96 | 1.42 | 51.4 |
| Bt54WS10 | 36.4 | 0.728 | 0.36 | 2.02 | 1.57 | 72.6 |
| BtF12 | 31.1 | 0.623 | 0.309 | 2.01 | 1.66 | 68.4 |
| BnWS09 | 21.1 | 0.421 | 0.204 | 2.06 | 1.37 | 42.4 |
| BnQD09 | 29.9 | 0.598 | 0.29 | 2.08 | 1.77 | 29.2 |
| BnQD12 | 29.5 | 0.589 | 0.285 | 2.06 | 1.67 | 50 |
| BjWS10 | 61 | 1.22 | 0.622 | 1.96 | 1.89 | 150 |
| BcWS09 | 38.1 | 0.761 | 0.371 | 2.05 | 1.26 | 63.8 |
| BcF09 | 19.1 | 0.382 | 0.182 | 2.09 | 1.41 | 35 |
| CcWA99 | 14.7 | 0.294 | 0.133 | 2.22 | 1.01 | 25.2 |
| DpF12 | 11 | 0.221 | 0.097 | 2.28 | 0.91 | 23.2 |

**Table C.3** Quality statistics for sequencing reads for 12 fruit fly samples.

| sample ID | Line | No. of flies | Barcode | No. of reads (pre-QC) | No. of reads (post-QC) | average sequence length (trimmed) | No. of reads (post chimera filtering) | Total % discarded |
|---|---|---|---|---|---|---|---|---|
| BtGWS09 | *B. tryoni* GOS | 8 | TCAGACGAGTGCGT | 4415 | 3874 | 428 | 3870 | 12.34 |
| BtHWS09 | *B. tryoni* HAC | 8 | TCAGACGCTCGACA | 9068 | 8672 | 429 | 8650 | 4.61 |
| Bt54WS10 | *B. tryoni* 54 | 8 | TCAGAGACGCACTC | 8789 | 8398 | 429 | 8384 | 4.61 |
| BtF12 | *B. tryoni* (field) | 8 | TCAGAGCACTGTAG | 8795 | 8404 | 428 | 8391 | 4.59 |
| BnWS09 | *B. neohumeralis* | 8 | TCAGATCAGACACG | 7191 | 6881 | 429 | 6873 | 4.42 |
| BnQD09 | *B. neohumeralis* 09 | 8 | TCAGATATCGCGAG | 9572 | 9137 | 427 | 9087 | 5.07 |
| BnQD12 | *B. neohumeralis* 12 | 8 | TCAGCGTGTCTCTA | 9424 | 8984 | 428 | 8734 | 7.32 |
| BjWS10 | *B. jarvisi* | 8 | TCAGCTCGCGTGTC | 18240 | 17514 | 428 | 17463 | 4.26 |
| BcWS09 | *B. cacuminata* | 8 | TCAGTACTGAGCTA | 11201 | 10700 | 429 | 10678 | 4.67 |
| BcF09 | *B. cacuminata* (field) | 8 | TCAGCATAGTAGTG | 10203 | 9748 | 429 | 9748 | 4.46 |
| CcWA99 | *C. capitata* Vienna7mix99 | 8 | TCAGTAGTATCAGC | 11533 | 11002 | 429 | 10998 | 4.64 |
| DpF12 | *D. pornia* (field) | 8 | TCAGTCTCTATGCG | 10937 | 10559 | 406 | 10281 | 6.00 |
| | | | **Total** | **119368** | **113873** | **426.6** | **113157** | |

**Table C.4** Jaccard distance matrix between samples, implemented in RDP II.

| | BtGWS09 | BtHWS09 | Bt54WS10 | BtF12 | BnWS09 | BnQD09 | BnQD12 | BjWS10 | BcWS09 | BcF09 | CcWA99 | DpF12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BtGWS09 | 0 | 0.0214434 | 0.0303548 | 0.1348838 | 0.1255519 | 0.4556549 | 0.3953688 | 0.7120028 | 0.7217418 | 0.3219664 | 0.9988186 | 0.972217 |
| BtHWS09 | 0.0214434 | 0 | 0.01118624 | 0.4587888 | 0.3004044 | 0.415102 | 0.4869828 | 0.4966255 | 0.552523 | 0.5104226 | 0.9988194 | 0.9755505 |
| Bt54WS10 | 0.0303548 | 0.0111862 | 0 | 0.2718542 | 0.0515544 | 0.5710979 | 0.5127065 | 0.4909574 | 0.5430583 | 0.5219089 | 0.9988195 | 0.9624503 |
| BtF12 | 0.1348838 | 0.4587888 | 0.27185416 | 0 | 0.2974337 | 0.5864443 | 0.3383658 | 0.9995241 | 0.9997698 | 0.2594482 | 0.8641243 | 0.9761822 |
| BnWS09 | 0.1255519 | 0.3004044 | 0.05155444 | 0.2974337 | 0 | 0.5978603 | 0.4596303 | 0.7638712 | 0.9769172 | 0.3888968 | 0.9988193 | 0.9771617 |
| BnQD09 | 0.4556549 | 0.415102 | 0.5710979 | 0.5864443 | 0.5978603 | 0 | 0.1008837 | 0.5951194 | 0.9468648 | 0.8285785 | 0.8873234 | 0.9780772 |
| BnQD12 | 0.3953688 | 0.4869828 | 0.5127065 | 0.3383658 | 0.4596303 | 0.1008837 | 0 | 0.9155412 | 0.9894635 | 0.3970341 | 0.4250978 | 0.9740311 |
| BjWS10 | 0.7120028 | 0.4966255 | 0.49095738 | 0.9995241 | 0.7638712 | 0.5951194 | 0.9155412 | 0 | 0.0861173 | 0.9960536 | 1 | 0.9984451 |
| BcWS09 | 0.7217418 | 0.552523 | 0.5430583 | 0.9997698 | 0.9769172 | 0.9468648 | 0.9894635 | 0.0861173 | 0 | 0.9978025 | 1 | 0.9984438 |
| BcF09 | 0.3219664 | 0.5104226 | 0.52190894 | 0.2594482 | 0.3888968 | 0.8285785 | 0.3970341 | 0.9960536 | 0.9978025 | 0 | 0.998818 | 0.9768809 |
| CcWA99 | 0.9988186 | 0.9988194 | 0.99881953 | 0.8641243 | 0.9988193 | 0.8873234 | 0.4250978 | 1 | 1 | 0.998818 | 0 | 0.9869683 |
| DpF12 | 0.972217 | 0.9755505 | 0.9624503 | 0.9761822 | 0.9771617 | 0.9780772 | 0.9740311 | 0.9984451 | 0.9984438 | 0.9768809 | 0.9869683 | 0 |



**Table C.5** UniFrac distance matrix between samples, unweighted distances below the diagonal, weighted distances above the diagonal.

| | BtGWS09 | BtHWS09 | Bt54WS10 | BtF12 | BnWS09 | BnQD09 | BnQD12 | BjWS10 | BcWS09 | BcF09 | CcWA99 | DpF12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| BtGWS09 | 0 | 0.0235216 | 0.11560706 | 0.0847313 | 0.1153069 | 0.2089916 | 0.059464 | 0.0864694 | 0.0857695 | 0.0832729 | 0.0894803 | 0.3309209 |
| BtHWS09 | 0.6023423 | 0 | 0.10259334 | 0.100963 | 0.1085131 | 0.1971037 | 0.0577768 | 0.0726135 | 0.0887055 | 0.0981945 | 0.1064447 | 0.3357428 |
| Bt54WS10 | 0.5427386 | 0.590116 | 0 | 0.1917285 | 0.0101772 | 0.1019625 | 0.1054457 | 0.0753793 | 0.1841766 | 0.1909022 | 0.1992998 | 0.3488373 |
| BtF12 | 0.5840371 | 0.6474317 | 0.61189234 | 0 | 0.1916751 | 0.2661921 | 0.1307095 | 0.1518705 | 0.0651917 | 0.0527945 | 0.0640631 | 0.3330865 |
| BnWS09 | 0.6182466 | 0.4211722 | 0.64215106 | 0.6684683 | 0 | 0.1063436 | 0.1117115 | 0.0831379 | 0.1830519 | 0.1866921 | 0.1948016 | 0.3490066 |
| BnQD09 | 0.5272009 | 0.6093099 | 0.54524857 | 0.4025047 | 0.661511 | 0 | 0.1524177 | 0.1469873 | 0.2666768 | 0.282612 | 0.2915163 | 0.3691876 |
| BnQD12 | 0.6543967 | 0.4091137 | 0.64163409 | 0.5886765 | 0.6115738 | 0.4755546 | 0 | 0.0701749 | 0.137785 | 0.1362722 | 0.1427319 | 0.3521576 |
| BjWS10 | 0.479394 | 0.5588397 | 0.5329441 | 0.6116908 | 0.6193745 | 0.5578051 | 0.6609886 | 0 | 0.1282552 | 0.1637773 | 0.1732287 | 0.3253241 |
| BcWS09 | 0.6585212 | 0.1992488 | 0.64781242 | 0.6955313 | 0.5322926 | 0.6649948 | 0.4904543 | 0.6019505 | 0 | 0.042679 | 0.0492017 | 0.3142944 |
| BcF09 | 0.6595825 | 0.6634232 | 0.66795537 | 0.6813748 | 0.7812386 | 0.6655987 | 0.7257627 | 0.6180314 | 0.6647237 | 0 | 0.0207681 | 0.3419743 |
| CcWA99 | 0.8330279 | 0.7183964 | 0.86447008 | 0.799595 | 0.6941923 | 0.81994 | 0.7190331 | 0.8627184 | 0.7247954 | 0.8669351 | 0 | 0.3503987 |
| DpF12 | 0.6708046 | 0.6556398 | 0.61226501 | 0.6793273 | 0.7480209 | 0.660595 | 0.6381401 | 0.7376312 | 0.6567306 | 0.7148123 | 0.8388119 | 0 |

# Appendix D

**Table D.1** Primer sequences used in this study (excluding qRT-PCR primer, see Table D.3).

| Locus (primer type) | Primer name | Sequence (5'-3') | Melting Temperature (°C) | Direction | Region | Reference |
|---|---|---|---|---|---|---|
| *transformer* | | | | | | |
| (degenerate) | tra-1d | CKRTTMGGAARGGTCCTYACGC | 58.6 | forward | exon1b | this study |
| | tra1f | GWTGACATCYMTCAACATGAAC | 51.6 | forward | exon1b | this study |
| | tra1erev | GATRTCRTCKCKTTGAAACARAGGC | 57.3 | reverse | exon2a | this study |
| | tra4 | GAGGAKRTTKMMAATCRAYGGCG | 57.3 | forward | exon2a | this study |
| | tra4rev | GTCGAWWTGTGACGYTCTTTKCGC | 58.9 | reverse | exon2a | this study |
| | tra-7a | KGTTATKATTTGCGGTTGCG | 53.4 | reverse | exon2b | this study |
| | tra-6 | GACMGMATTCGACGTAAATATGG | 53.8 | forward | exon2b | this study |
| | tra-7b | AATCCGCWGGAACYGGCACKGG | 64.5 | reverse | exon2b | this study |
| | tra-8 | CTWWAWCCASGTCCACCG | 52.7 | reverse | exon3 | this study |
| (*Bactrocera* specific) | traex1for | TTGAACGCCATCGCGACAAAC | 58.9 | forward | exon1a | this study |
| | tracDNAL2 | TATCTATTGAACGCCATCGC | 52.6 | forward | exon1a | this study |
| | tracDNAL1a | TATCTATTGAACGCCATCGCGAC | 57.1 | forward | exon1a | this study |
| | traex2 | AGTTCCAGACGAAGTTGTTATTAAGC | 55.3 | forward | exon1b | this study |
| | traex3 | AGATGATATCGTGGTGAATCCG | 54.3 | forward | exon2a | this study |
| | traex3rev | GCCCTCGATTTGAAATATCCTCTG | 55.6 | reverse | exon2a | this study |
| | tra6brev | GCTAGAGGCGATTTATTTCTTGTCG | 55.9 | reverse | exon2b | this study |
| | tra6drev | CTCGAGAGTGAGATCGTGAGCGTG | 60.7 | reverse | exon2b | this study |
| | tra8c | CCCACTTATCCTATGCCTACTTTCG | 56.8 | forward | exon3 | this study |
| | tra8b | GGATTACCTCCACAACCAATACG | 55.4 | forward | exon3 | this study |
| | tracDNAR2 | AAGTGTATTTTTAAATTTACTTGCTTTTTG | 51.3 | reverse | exon3 | this study |

| Locus (primer type) | Primer name | Sequence (5'-3') | Melting Temperature (°C) | Direction | Region | Reference |
|---|---|---|---|---|---|---|
| *transformer-2* | | | | | | |
| (degenerate) | 2TRA-A | GCAAGCGGGCATTTWYATTTYAATC | 55.8 | forward | exon1 | this study |
| | 2TRA-D | CACGACAMCCGCCTTCWCCACC | 63.7 | forward | exon3 | this study |
| | 2TRA-E | CAGAWGCCAGCCAAAGYTCTTCAAC | 59.6 | forward | exon3 | this study |
| | 2TRA-F | CAGTRCAAAAYCGTTGTATAGG | 51.6 | forward | exon5 | this study |
| | 2TRA-Z | GTTGCGTTGTATAMACACTYAAACC | 54.8 | reverse | exon5 | this study |
| | 2TRA-H | GACCMATCGARMGRATACAAGTC | 55.3 | forward | | this study |
| | 2TRA-X | CAAGYGTCTYTRGCGCRCYACTC | 61.2 | reverse | intron5 | this study |
| | 2TRA-Q | GAGAAYGKGAACGKGWYATMCGACGG | 61.7 | reverse | exon4 | this study |
| | 2TRA-P | GTGAC/GTTGAAGARCTTTGGCTG | 57.7 | reverse | exon3/4 | this study |
| | 2TRA-S | TKAWARGGTGARCCYGARCGACGG | 61.8 | reverse | exon6b | this study |
| (*Bactrocera* specific) | 2traEs | CAGAAGCCAGCCAAAGCTCTTCAAC | 60.5 | forward | exon3 | this study |
| | 2tra-p | CGACGCTTTTCGTATGTGCTAC | 56.3 | reverse | exon4 | this study |
| | 2tra-i | CAGTGCAAAATCGGTGTATAGG | 53.8 | forward | exon5 | this study |
| | 2traeq1 | AATTTTAACGGGCGCAGGCATACGC | 62.5 | | | this study |
| | 2traeq1rev | AATATAAGCGACGGACCGTGCG | 59.6 | | | this study |
| | 2trahs1 | CGTTTATATGGGACGACACAC | 53.3 | | | this study |
| | 2trahs1rev | TACTCACCTGTGCATCAATGACG | 57.1 | | | this study |
| | 2traex122 | GTAGTTATAATG/AGCCCTCGTTCACG | 57 | forward | exon1/2 | this study |
| | 2traex425 | ACTGGCTTTTC/CCTATTACCACGAC | 58.8 | reverse | exon4/5 | this study |
| | 2tracDNAL1 | CATTATTGTGATTTGTCAGCTC | 50 | forward | exon1 | this study |
| | 2tracDNAR1 | TAAGTGAGCAAATAATAAATGGTG | 49.5 | reverse | exon7 | this study |
| Tra-2 pseudogene | | | | | | |
| | BjY2traA | GAATAGTTATAATG/AGCCCTCGTTCACG | 57.1 | forward | exon1/2 | D.C.A. Shearman |
| | BjY2traB | AATGGGCGTCG/ACGATATTCAAAAG | 57.7 | forward | exon2/3 | D.C.A. Shearman |
| | BjY2traDrev | CACTGGCTTTTC/CCTATTACCACGAC | 59.4 | reverse | exon4/5 | D.C.A. Shearman |
| | BjY2traErev | CACGTTCACGC/CGCGTTTGTCGG | 65.4 | reverse | exon6a/6b | D.C.A. Shearman |

| Locus (primer type) | Primer name | Sequence (5'-3') | Melting Temperature (°C) | Direction | Region | Reference |
|---|---|---|---|---|---|---|
| *Sex-lethal* | | | | | | |
| | SXL1 | TGTATGGGAATATGAATAATGG | 47.7 | forward | | D.C.A. Shearman |
| | SXL2 | CACAATTCAGCGAATTTGTGC | 53.4 | reverse | | D.C.A. Shearman |
| | SXL3 | ATGGATACAGATTTCACCTCATC | 52.2 | forward | | D.C.A. Shearman |
| | SXL4 | TTATCGCTCGTTGTGAATCC | 52.7 | reverse | | D.C.A. Shearman |
| | SXL5 | CCGAAACGGATTCACAACGAGC | 58.8 | forward | | D.C.A. Shearman |
| | SXL6 | TAGTGTTGCGCTTTCGCCTTG | 58.5 | reverse | | D.C.A. Shearman |
| | SXL7 | CTCTGAACAATGTCATACCCGAG | 55.1 | forward | | D.C.A. Shearman |
| | SXL8 | ATGAATTTCTGTGCATTATGATAGG | 51.6 | reverse | | D.C.A. Shearman |
| | SxlRTfor1 | AGACAAATTGACGGGCAAACCACG | 60.4 | forward | | this study |
| | SxlRTrev1 | GACATGTATTGCTGCGCTTTCGCT | 60.4 | reverse | | this study |
| 3' RACE | | | | | | |
| 1st strand cDNA synthesis | oligo (dT) | | | | | |
| | polyT adapter | GTGACTCGAGTCGACATCGA (T)17 | 59.1 | | | |
| | adapter | GTGACTCGAGTCGACATCG | 55 | | | |
| *tra*-specific | tra8c | CCCACTTATCCTATGCCTACTTTCG | 56.8 | forward | exon3 | this study |
| | tra8b | GGATTACCTCCACAACCAATACG | 55.4 | forward | exon3 | this study |
| *tra-2*-specific | 2traEs | CAGAAGCCAGCCAAAGCTCTTCAAC | 60.5 | forward | exon3 | this study |
| | 2tra-i | CAGTGCAAAATCGGTGTATAGG | 53.8 | forward | exon5 | this study |
| 5' RACE | | | | | | |
| 1st strand cDNA synthesis | traex3rev | GCCCTCGATTTGAAATATCCTCTG | 55.6 | reverse | exon2a | this study |
| | 2tra-p | CGACGCTTTTCGTATGTGCTAC | 56.3 | reverse | exon4 | this study |
| | AAP | GGCCACGCGTCGACTAGTAC (G)14 | 77 | | | |
| | AUAP | GGCCACGCGTCGACTAGTAC | 60.1 | | | |
| *tra*-specific | tra5race1 | CTACGCTTAATAACAACTTCGTCTGG | 55.6 | reverse | | D.C.A. Shearman |
| | tra5race2 | GTGAGGACCCTTCCTAATTGAGCC | 59.2 | reverse | | D.C.A. Shearman |
| *tra-2*-specific | 2tra5race1 | AGTAACGTCCACCTCGTCCTGATGGC | 63.7 | reverse | | D.C.A. Shearman |
| | 2tra5race2 | GTTGAAGAGCTTTGACTGGCTTCTGGC | 61.8 | reverse | | D.C.A. Shearman |

**Table D.2** PCR conditions and primer sequences used in this research. Degenerate PCR protocol was applied to genomic DNA with combinations of forward and reverse primers for *tra* and *tra-2*. RT-PCR protocol was applied to cDNA with gene-specific primers. 3' and 5' RACE were nested reactions using adapter primers and nested gene-specific primers. Y-chromosome and Multiplex Y-chromosome marker protocols were both used for non-sex-specific *Sxl* and sex-specific Y primers, as single primer pair reactions or multiplex as needed. Primer sequences are listed in Table D.1.

| **PCR protocol** | **Degenerate PCR** | **RT-PCR** | **3' and 5' RACE** | **Y-Chromosome marker** | **Multiplex Y-Chromosome** |
|---|---|---|---|---|---|
| Total volume | 25µL | 20µL | 20µL | 10µL | 10µL |
| GoTaq Reaction Buffer | 1X | 1X | 1X | 1X | 1X |
| MgCl2 | 1.5mM | 2.5mM | 1.5mM | 2.5mM | 2.5mM |
| dNTP | 200µM each | 200µM each | 250µM each | 200µM each | 200µM each |
| Primer 1 | 0.5µM | 0.5µM | 0.4µM | 0.375µM | 0.375µM |
| Primer 2 | 0.5µM | 0.5µM | 0.4µM | 0.375µM | 0.375µM |
| Primer 3 | | | | | 0.375µM |
| Primer 4 | | | | | 0.375µM |
| GoTaq DNA polymerase | 1.25u | 0.5u | 1u | 0.5u | 0.5u |
| DNA/cDNA Template | 1µL | 0.5µL | 3µL | 2µL | 2µL |
| **Thermal Cycling** | | | | | |
| Denaturation | 94°C for 3min | 94°C for 3min | 94°C for 2min | 94°C for 5min | 94°C for 5min |
| Cycling | 94°C for 1min, 55°C for 1min, 72°C for 1min (x30cycles) | 94°C for 30s, 60°C for 30s, 72°C for 1min (x30cycles) | 94°C for 30s, 55°C for 30s, 72°C for 80s (x30cycles) | 94°C for 30s, 60°C for 30s, 72°C for 1min (x35 cycles) | 94°C for 30s, 60°C for 30s, 72°C for 1min (x35 cycles) |
| Final extension | 72°C for 5min | 72°C for 5min | 72°C for 7min | 72°C for 7min | 72°C for 7min |

PCR reagents are from Promega (GoTaq DNA polymerase), Bioline (dNTPs), Macrogen (oligonucleotides)

Thermocycler is the BioRad DNA Engine Dyad (BioRad)

**Table D.3** qRT-PCR primers used in Chapter 5.

| Species | Locus | qPCR primers | Sequence (5'-3') | Melting temp (°C)* | Region | product size | Reference |
|---|---|---|---|---|---|---|---|
| BJ and BTJ[Y] | *Sxl* | SxlRTfor2 | GCATTTGTCAG/ATTCAATAAGCGCGA | 58.2 | | 143bp | this study |
| | | SxlRTrev1 | GACATGTATTGCTGCGCTTTCGCT | 60.4 | | | this study |
| BJ and BTJ[Y] | *dsxF* | dsxFRTfor1 | ACCATCTCTGCATCCCAATGGAGT | 60.2 | exon2 | 189bp | this study |
| | | dsxFRTrev2 | CGTTTACGACATGTTGGC/CTTCCT | 59.6 | exon3/4 | | this study |
| BJ and BTJ[Y] | *dsxM* | dsxMRTfor1 | GACGCATTGAGGAAG/CAAAGCGAA | 60.3 | exon 3/5 | 180bp | this study |
| | | dsxMRTrev1 | TTAGGTGCGAGAAGTGCGAAGTCA | 60.3 | exon5 | | this study |
| BJ and BTJ[Y] | *tra-2* | 2traBtRTfor1 | CAAACTCG/GCGTGAACGTGAGCAT | 62 | exon6a/6b | 148bp | this study |
| | | 2traBtRTrev1 | AGTAGGAACGACTCCGGCTGCGTT | 64.1 | exon6b | | this study |
| BJ and BTJ[Y] | *traM* | traBtMRTfor1 | TGAGACAAGGTGTTAGCTTGC | 55.2 | MS1 | 136bp | this study |
| | | traBtMRTrev1 | CGGCTGCTTCTAAAGG/TTTATT | 53.3 | MS1/MS2 | | this study |
| BTJ[Y] | *traF* | traBtFRTfor1 | AGCGTAGATTCG/GTGAAGGT | 55.9 | exon1b/2a | 120bp | this study |
| | | traBtFRTrev1 | TATCCTCTGTGGTGCTTTGC | 55 | exon 2a | | this study |
| BJ | *traF* | traBtFRTfor1 | AGCGTAGATTCG/GTGAAGGT | 55.9 | exon1b/2a | 183bp | this study |
| | | traBjFRTrev1 | TGGAATC/AGTCGAAATGTGG | 52.7 | exon 2a/2b | | this study |
| BJ and BTJ[Y] | *slam* | BtslamFor1 | AGGGACGTGGTCGTACCGCT | 63.1 | | 155bp | this study |
| | | BtslamRev1 | GGTGCCGGTTGGCGATTCA | 61.2 | | | this study |
| BJ and BTJ[Y] | *leo* | LeoFor | GTTTTCTGGCGGAGGATTCAG | 56.1 | | 189bp | Cook, 2005 |
| | | LeoRev | AGATACGATGACATGGCCCAG | 56.7 | | | Cook, 2005 |
| Exogenous | *Kan* | K1 | ATCAGGTGCGACAATCTATCG | 54.6 | | 763bp | An, et. al., 2002 |
| | | K2 | CGTCAAGTCAGCGTAATGCTC | 56.1 | | | An, et. al., 2002 |

*Melting temperatures were calculated by OligoAnalyzer (http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/)

**Table D.4** qRT-PCR efficiency values for each primer pair. Each qRT-PCR primer pair (Table D.3) was tested on a serial dilution of a purified amplicon. Threshold level was arbitrarily set and RotorGene 6000 software produced the efficiency curves.

| **Locus** | **R** | **$R^2$** | **M** | **B** | **Efficiency** | **Threshold** |
|---|---|---|---|---|---|---|
| *Sex Lethal* | 0.994 | 0.989 | -3.32 | 2.21 | 1 | 0.0555 |
| *doublesex* female | 0.996 | 0.993 | -3.49 | 2.80 | 0.96 | 0.0555 |
| *doublesex* male | 0.998 | 0.995 | -3.30 | 1.22 | 1.01 | 0.0555 |
| *transformer-2* | 0.998 | 0.996 | -3.45 | 2.99 | 0.95 | 0.0555 |
| *transformer* female | 0.996 | 0.991 | -3.23 | 2.94 | 1.04 | 0.0555 |
| *transformer* male | 0.999 | 0.998 | -3.24 | 4.19 | 1.03 | 0.0555 |
| *slow as molasses* | 0.999 | 0.999 | -3.16 | 4.05 | 1.07 | 0.0555 |
| *leonardo* | 0.999 | 0.999 | -3.26 | -0.94 | 1.03 | 0.0555 |
| *Kanamycin* | 0.992 | 0.984 | -3.68 | -1.59 | 0.87 | 0.0555 |
| Alien inhibitor | 0.998 | 0.996 | -3.24 | 1.10 | 1.03 | 0.0555 |

# Appendix E

**Table E.1**. PCR conditions and primer sequences used in Chapter 6.

| PCR protocol | Y-Chromosome marker | Multiplex Y-Chromosome |
|---|---|---|
| Reaction volume | 10µL | 10µL |
| GoTaq Reaction Buffer | 1X | 1X |
| $MgCl_2$ | 2.5mM | 2.5mM |
| dNTP | 200µM each | 200µM each |
| Primer 1 | 0.375µM | 0.375µM |
| Primer 2 | 0.375µM | 0.375µM |
| Primer 3 | | 0.375µM |
| Primer 4 | | 0.375µM |
| GoTaq DNA polymerase | 0.5u | 0.5u |
| DNA Template | 2µL | 2µL |
| | | |
| **Thermal Cycling** | 94°C for 5min | 94°C for 5min |
| | 94°C for 30s, 60°C for 30s, 72°C for 1min (x35 cycles) | 94°C for 30s, 60°C for 30s, 72°C for 1min (x35 cycles) |
| | 72°C for 7min | 72°C for 7min |
| **Primers** | **Sequence 5'-3'** | **Combinations (product size)** |
| BjY2traA | GAATAGTTATAATG/AGCCCTCGTTCACG | BjY2traA-BjY2traDrev (311bp) |
| BjY2traB | AATGGGCGTCG/ACGATATTCAAAAG | BjY2traB-BjY2traDrev (227bp) |
| BjY2traDrev | CACTGGCTTTTC/CCTATTACCACGAC | |
| SxlRTfor1 | AGACAAATTGACGGGCAAACCACG | SxlRTfor1 - SxlRTrev1 (280bp) |
| SxlRTrev1 | GACATGTATTGCTGCGCTTTCGCT | |

PCR reagents are from Promega (GoTaq DNA polymerase), Bioline (dNTPs), Macrogen (oligonucleotides); thermocycler is the BioRad DNA Engine Dyad.

**Table E.2**. Quality and concentration of total RNA samples for transcriptome sequencing.

| | | | No. of | Nanodrop | | | Qubit |
|---|---|---|---|---|---|---|---|
| **Sample ID** | **sex** | **age (AEL)** | **embryos** | **ng/uL** | **260/280nm** | **260/230nm** | **ng/uL** |
| BJ1 | male | 3-5h | 26 | 103.8 | 1.99 | 2.02 | 63 |
| BJ2 | male | 3-5h | 23 | 177.5 | 2.07 | 1.74 | 115 |
| BJ3 | female | 3-5h | 27 | 193.3 | 2.09 | 1.65 | 119 |
| BJ4 | female | 3-5h | 17 | 137.5 | 2.05 | 1.82 | 93 |
| BJ5 | male | 2-3h | 28 | 120.6 | 2.07 | 2.04 | 84 |
| BJ6 | male | 2-3h | 24 | 208.1 | 2 | 1.77 | 125 |
| BJ7 | female | 2-3h | 24 | 268 | 2.09 | 1.76 | 104 |
| BJ8 | female | 2-3h | 12 | 101.3 | 2.03 | 1.8 | 59 |

**Table E.3** Sex-determination and cellularisation genes. **(A)** The 28 *Drosophila* sex determination genes and the top BLAST hits against the nr database. Only 19 of these genes found a match in the CLC assembly. **(B)** Four cellularisation genes were sought, only nullo was found by the top blast hits. Other transcripts were identified by motif searches using *D. melanogaster* or *C. capitata* sequence as the query.

| A | Annotation Symbol | contig no. | contig size | E-value | Top blast hit | Description |
|---|---|---|---|---|---|---|
| *daughterless* | CG5102 | 16154 | 4053 | 0 | XP_004521529 | PREDICTED: protein daughterless-like [Ceratitis capitata] |
| *deadpan* | CG8704 | 20953 | 3365 | 0 | XP_004534624 | PREDICTED: protein deadpan-like [Ceratitis capitata] |
| *degringolade* | | | | | | no match |
| *dissatisfaction* | | | | | | no match |
| *doublesex* | CG11094 | 22304 | 1064 | 2.84E-72 | ACN24617 | female-specific doublesex protein [Bactrocera dorsalis] |
| | | 40032 | 808 | 1.35E-58 | AAB99947 | doublesex [Bactrocera tryoni] |
| *extra-macrochaetae* | CG1007 | 10611 | 1230 | 2.02E-67 | XP_004525703 | PREDICTED: protein extra-macrochaetae-like [Ceratitis capitata] |
| *female lethal d* | CG6315 | 2854 | 2694 | 0 | XP_004520275 | PREDICTED: pre-mRNA-splicing regulator female-lethal(2)D-like isoform X3 [Ceratitis capitata] |
| *fruitless* | | 4880 | 1862 | 1.14E-75 | XP_004523736 | PREDICTED: sex determination protein fruitless-like [Ceratitis capitata] |
| | | 29049 | 3175 | 4.28E-104 | AGL09914 | male-specific zinc finger C splice variant [Musca domestica] |
| *groucho* | CG8384 | 3466 | 4604 | 0 | XP_004518043 | PREDICTED: protein groucho-like isoform X1 [Ceratitis capitata] |
| *hermaphrodite* | | | | | | no match |
| *hopscotch* | CG1594 | 753 | 4443 | 0 | XP_004518856 | PREDICTED: tyrosine-protein kinase hopscotch-like [Ceratitis capitata] |
| *intersex* | | | | | | no match |
| *Mes-4* | CG4976 | 5091 | 5397 | 0 | XP_004530722 | PREDICTED: probable histone-lysine N-methyltransferase Mes-4-like isoform X1 [Ceratitis capitata] |
| *Mutant-181* | | | | | | no match |
| *outstretched* | | | | | | no match |
| *ovarian tumor* | CG12743 | 3340 | 3198 | 1.41E-72 | XP_004527434 | PREDICTED: protein ovarian tumor locus-like isoform X3 [Ceratitis capitata] |
| | | 5442 | 201 | 7.48E-09 | XP_004527433 | PREDICTED: protein ovarian tumor locus-like isoform X2 [Ceratitis capitata] |
| | | 5443 | 201 | 1.44E-08 | XP_004527433 | PREDICTED: protein ovarian tumor locus-like isoform X2 [Ceratitis capitata] |
| | | 8947 | 1757 | 0 | XP_004527434 | PREDICTED: protein ovarian tumor locus-like isoform X3 [Ceratitis capitata] |
| | | 10161 | 300 | 5.23E-24 | XP_004527434 | PREDICTED: protein ovarian tumor locus-like isoform X3 [Ceratitis capitata] |
| *ovo* | CG6824 | 21175 | 255 | 5.10E-26 | XP_004527388 | PREDICTED: protein ovo-like isoform X1 [Ceratitis capitata] |
| | | 22105 | 317 | 3.13E-12 | XP_004527393 | PREDICTED: protein ovo-like isoform X6 [Ceratitis capitata] |
| | | 23602 | 317 | 3.13E-12 | XP_004527393 | PREDICTED: protein ovo-like isoform X6 [Ceratitis capitata] |
| *PHD finger protein 7 ortholog* | | | | | | no match |

| A (cont) | Annotation Symbol | contig no. | contig size | E-value | Top blast hit | Description |
|---|---|---|---|---|---|---|
| *runt* | CG1849 | 13344 | 1521 | 2.57E-136 | XP_004527365 | PREDICTED: segmentation protein Runt-like isoform X1 [Ceratitis capitata] |
| | | 1493 | 840 | 1.33E-26 | | |
| | | 1494 | 229 | 9.87E-27 | | |
| *sans fille* | CG4528 | 7153 | 1421 | 4.99E-91 | P43332 | RecName: Full=U1 small nuclear ribonucleoprotein A; Short=U1 snRNP A; Short=U1-A; Short=U1A; AltName: Full=Sex determination protein snf |
| *scute* | CG3827 | 57147 | 340 | 5.11E-57 | AAF66944 | scute [Ceratitis capitata](389aa) |
| | | 50842 | 514 | 1.79E-10 | AAF66944 | scute [Ceratitis capitata] |
| *Sex-lethal* | CG43770 | 2265 | 1667 | 1.92E-140 | CAG29242 | sex-lethal protein [Bactrocera oleae] |
| *Stat92E* | | | | | | no match |
| *sisterless A* | CG1641 | 2508 | 264 | 4.66E-09 | XP_002055743 | sisA [Drosophila virilis] >gi\|194150253\|gb\|EDW65944.1\| |
| | | | | | XM_004527132 | PREDICTED: Ceratitis capitata protein sisterless A-like (LOC101450422) mRNA |
| *stand still* | | | | | | no match |
| *transformer* | CG16724 | 4523 | 3261 | 1.93E-150 | XP_004526946 | PREDICTED: polycomb group RING finger protein 3-like [Ceratitis capitata] |
| | | 13992 | 392 | 4.30E-23 | AAZ08052 | transformer male-specific 1 [Bactrocera oleae] |
| *transformer 2* | CG10128 | 138 | 3165 | 6.62E-39 | O02008 | RecName: Full=Transformer-2 sex-determining protein |
| *virilizer* | CG3496 | 2897 | 994 | 1.04E-148 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 4042 | 392 | 3.41E-51 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 4266 | 3049 | 0 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 4267 | 1131 | 1.42E-120 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 5176 | 234 | 2.77E-30 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 5177 | 234 | 1.71E-30 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 7162 | 268 | 9.10E-32 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 7163 | 268 | 1.56E-30 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 12608 | 218 | 1.72E-12 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 16540 | 309 | 2.04E-04 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |
| | | 16541 | 310 | 1.25E-08 | XP_004525965 | PREDICTED: LOW QUALITY PROTEIN: protein virilizer-like [Ceratitis capitata] |

<table>
<tr><th>B</th><th>Annotation Symbol</th><th>contig no.</th><th>contig size</th><th>E-value</th><th>Top blast hit</th><th>Description</th></tr>
<tr><td>bottleneck</td><td></td><td></td><td></td><td></td><td></td><td>no match</td></tr>
<tr><td>nullo</td><td>CG14462</td><td>9415</td><td>202</td><td>0.18</td><td>AAB46422</td><td>nullo, partial [Drosophila simulans]</td></tr>
<tr><td></td><td></td><td>6453</td><td>764</td><td>7.82E-06</td><td>XP_001978515</td><td>GG19631 [Drosophila erecta] >gi|190650164|gb|EDV47442.1|</td></tr>
<tr><td></td><td></td><td>14101</td><td>1026</td><td></td><td></td><td>no match</td></tr>
<tr><td></td><td></td><td>14102</td><td>593</td><td></td><td></td><td>no match</td></tr>
<tr><td></td><td></td><td>14117</td><td>1443</td><td></td><td></td><td>no match</td></tr>
<tr><td></td><td></td><td>15641</td><td>455</td><td></td><td></td><td>no match</td></tr>
<tr><td>serendipity α</td><td>CG17957</td><td>9105</td><td>5007</td><td>0</td><td>XP_004536849</td><td>PREDICTED: GPI mannosyltransferase 4-like [Ceratitis capitata]</td></tr>
<tr><td>slow as molasses</td><td>CG9506</td><td>3824</td><td>5366</td><td>0</td><td>XP_004534436</td><td>PREDICTED: titin-like [Ceratitis capitata]</td></tr>
</table>

**Table E.4.** A list of the 66 contigs up-regulated in older male embryos, after exclusion of contigs also up-regulated in older females. Normalised RPKM values are from experiment I (Table 6.6) with top BLAST hits to arthropod sequences from the NCBI nr database.

| | | | Normalised RPKM values | | | | | | | | Top BLAST Hit | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Contig ID | Contig length | Fold Change | male BJ2 | female BJ3 | female BJ4 | “late” mean | male BJ6 | female BJ7 | female BJ8 | “early” mean | Accession (E-value) | Description |
| 36 | 665 | 28.54 | 88.1 | 80.1 | 45.2 | 71.1 | 0.7 | 0.0 | 6.8 | 2.5 | XP_004525671 | PREDICTED: uncharacterized protein LOC101459031 [Ceratitis capitata] |
| 645 | 589 | 21.01 | 25.8 | 24.7 | 13.3 | 21.3 | 0.3 | 0.1 | 2.6 | 1.0 | ABD76335 | transposase [Heliothis virescens] |
| 930 | 564 | 4.27 | 62.1 | 55.1 | 78.8 | 65.3 | 11.7 | 17.6 | 16.6 | 15.3 | XP_004521270 | PREDICTED: eukaryotic translation initiation factor 3 subunit D-1-like [Ceratitis capitata] |
| 1038 | 1089 | 3.54 | 56.1 | 52.4 | 78.0 | 62.2 | 13.2 | 22.6 | 16.9 | 17.6 | XP_004521155 | PREDICTED: uncharacterized protein LOC101455781 [Ceratitis capitata] |
| 1228 | 2739 | 4.06 | 198.5 | 201.5 | 131.6 | 177.2 | 37.6 | 45.7 | 47.7 | 43.7 | XP_004521206 | PREDICTED: eukaryotic initiation factor 4A-III-like [Ceratitis capitata] |
| 1817 | 1774 | 7.33 | 75.7 | 75.5 | 129.4 | 93.5 | 14.1 | 10.8 | 13.4 | 12.8 | XP_004537461 | PREDICTED: craniofacial development protein 1-like [Ceratitis capitata] |
| 2314 | 4147 | 5.98 | 156.6 | 162.2 | 265.5 | 194.8 | 23.1 | 38.9 | 35.9 | 32.6 | XP_004523499 | PREDICTED: DNA primase large subunit-like [Ceratitis capitata] |
| 2876 | 3809 | 2.23 | 115.8 | 103.0 | 79.9 | 99.6 | 48.6 | 37.3 | 47.9 | 44.6 | XP_004521895 | PREDICTED: transportin-3-like [Ceratitis capitata] |
| 3007 | 223 | 74.82 | 54.8 | 66.4 | 37.3 | 52.8 | 0.8 | 0.0 | 1.3 | 0.7 | AAB82058 | blastoderm-specific protein 25A [Drosophila melanogaster] |
| 3009 | 1230 | 13.81 | 28.5 | 26.5 | 14.5 | 23.2 | 1.9 | 0.3 | 2.8 | 1.7 | XP_001988422 | >gi\|193904422\|gb\|EDW03289.1\| GH10578 [Drosophila grimshawi] |
| 3156 | 509 | 4.16 | 40.1 | 35.9 | 60.7 | 45.6 | 9.3 | 12.6 | 11.0 | 11.0 | XP_004531652 | PREDICTED: calnexin-like [Ceratitis capitata] |
| 3353 | 982 | 11.51 | 88.3 | 76.9 | 44.7 | 70.0 | 4.6 | 0.3 | 13.3 | 6.1 | XP_002066566 | >gi\|194162651\|gb\|EDW77552.1\| GK24494 [Drosophila willistoni] |
| 3466 | 4604 | 6.37 | 51.1 | 59.2 | 95.9 | 68.8 | 7.4 | 13.0 | 12.0 | 10.8 | XP_004518043 | PREDICTED: protein groucho-like isoform X1 [Ceratitis capitata] |
| 3515 | 2772 | 4.26 | 88.8 | 92.8 | 135.3 | 105.6 | 23.1 | 25.1 | 26.2 | 24.8 | XP_004529388 | PREDICTED: keratin, type I cytoskeletal 9-like [Ceratitis capitata] |
| 3878 | 3339 | 46.38 | 73.1 | 85.2 | 46.1 | 68.1 | 1.2 | 1.1 | 2.2 | 1.5 | XP_004521898 | PREDICTED: TNF receptor-associated factor 4-like isoform X1 [Ceratitis capitata] |
| 4660 | 4268 | 14.5 | 133.7 | 134.4 | 213.3 | 160.5 | 5.9 | 14.1 | 13.2 | 11.1 | XP_004529208 | PREDICTED: histone-arginine methyltransferase CARMER-like [Ceratitis capitata] |
| 4679 | 1184 | 13.79 | 256.9 | 192.6 | 126.5 | 192.0 | 7.0 | 0.6 | 34.2 | 13.9 | XP_004522978 | PREDICTED: poly(A) polymerase alpha-like [Ceratitis capitata] |
| 4783 | 326 | 15.69 | 261.6 | 263.9 | 165.5 | 230.4 | 3.6 | 15.1 | 25.3 | 14.7 | XP_004517515 | PREDICTED: F-box/LRR-repeat protein 14-like isoform X2 [Ceratitis capitata] |
| 4784 | 2318 | 10.95 | 409.6 | 336.4 | 257.4 | 334.5 | 12.6 | 34.2 | 44.9 | 30.6 | XP_004517511 | PREDICTED: F-box/LRR-repeat protein 13-like [Ceratitis capitata] |
| 4801 | 3454 | 3.65 | 123.3 | 124.4 | 87.4 | 111.7 | 25.2 | 27.6 | 39.0 | 30.6 | XP_004534997 | PREDICTED: polymerase delta-interacting protein 3-like [Ceratitis capitata] |
| 4934 | 2359 | 14.06 | 163.4 | 160.5 | 96.1 | 140.0 | 7.4 | 7.2 | 15.3 | 10.0 | XP_004535595 | PREDICTED: parafibromin-like [Ceratitis capitata] |
| 4984 | 2762 | 20.4 | 320.0 | 302.8 | 176.1 | 266.3 | 8.0 | 8.6 | 22.6 | 13.1 | XP_001988422 | >gi\|193904422\|gb\|EDW03289.1\| GH10578 [Drosophila grimshawi] |
| 5000 | 665 | 9.25 | 66.8 | 50.0 | 33.7 | 50.2 | 1.2 | 0.2 | 14.9 | 5.4 | XP_004524168 | PREDICTED: protein toll-like [Ceratitis capitata] |

| | | | Normalised RPKM values | | | | | | | | Top BLAST Hit | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Contig ID | Contig length | Fold Change | male BJ2 | female BJ3 | BJ4 | “late” mean | male BJ6 | female BJ7 | BJ8 | “early” mean | Accession (E-value) | Description |
| 6083 | 406 | 44.33 | 139.2 | 150.8 | 85.2 | 125.1 | 1.9 | 1.0 | 5.5 | 2.8 | AAB82058 | blastoderm-specific protein 25A [Drosophila melanogaster] |
| 6152 | 629 | 8.26 | 183.9 | 153.9 | 93.4 | 143.7 | 7.8 | 15.7 | 28.8 | 17.4 | XP_004518482 | PREDICTED: serine/arginine repetitive matrix protein 2-like [Ceratitis capitata] |
| 6356 | 793 | 16.87 | 48.5 | 54.6 | 32.4 | 45.2 | 2.4 | 0.1 | 5.5 | 2.7 | XP_001355983 | >gi\|54644301\|gb\|EAL33042.1\| GA22097 [Drosophila pseudoobscura pseudoobscura] |
| 6406 | 256 | 14.23 | 35.9 | 26.3 | 16.6 | 26.3 | 2.4 | 0.4 | 2.7 | 1.9 | XP_004519076 | PREDICTED: uncharacterized protein LOC101456548 [Ceratitis capitata] |
| 6564 | 2000 | 17.45 | 162.9 | 137.6 | 81.6 | 127.4 | 1.6 | 5.4 | 14.9 | 7.3 | XP_004537817 | PREDICTED: mediator of RNA polymerase II transcription subunit 15-like [Ceratitis capitata] |
| 7120 | 2686 | 51.74 | 22.1 | 19.6 | 10.0 | 17.2 | 0.2 | 0.1 | 0.7 | 0.3 | XP_004530768 | PREDICTED: uncharacterized protein LOC101463003 [Ceratitis capitata] |
| 7485 | 1504 | 8.07 | 79.5 | 70.8 | 42.6 | 64.3 | 3.4 | 8.0 | 12.6 | 8.0 | XP_004522803 | PREDICTED: ATP-dependent RNA helicase DHX8-like isoform X2 [Ceratitis capitata] |
| 7654 | 216 | 3.9 | 65.5 | 59.7 | 86.3 | 70.5 | 14.9 | 13.5 | 25.9 | 18.1 | XP_004517756 | PREDICTED: eukaryotic translation initiation factor 3 subunit B-like [Ceratitis capitata] |
| 7816 | 201 | 3.88 | 33.6 | 24.7 | 31.8 | 30.0 | 4.9 | 6.5 | 11.8 | 7.7 | XP_004521644 | PREDICTED: protein BUD31 homolog [Ceratitis capitata] |
| 8666 | 3357 | 8.49 | 82.9 | 82.8 | 137.3 | 101.0 | 6.8 | 15.6 | 13.3 | 11.9 | XP_004534405 | PREDICTED: protein ref(2)P-like [Ceratitis capitata] |
| 8873 | 431 | 3.32 | 29.7 | 28.5 | 23.6 | 27.3 | 4.5 | 8.5 | 11.7 | 8.2 | XP_004520278 | PREDICTED: pre-mRNA-splicing factor SYF1-like [Ceratitis capitata] |
| 8897 | 2006 | 17.81 | 38.9 | 36.1 | 18.0 | 31.0 | 0.3 | 0.1 | 4.8 | 1.7 | ACO12003 | Transposable element Tc3 transposase [Lepeophtheirus salmonis] |
| 8928 | 881 | 2.9 | 107.5 | 91.4 | 121.2 | 106.7 | 25.5 | 41.6 | 43.2 | 36.8 | XP_004521190 | PREDICTED: mitotic checkpoint protein BUB3-like [Ceratitis capitata] |
| 10611 | 1230 | 125.04 | 131.1 | 131.6 | 225.0 | 162.6 | 0.6 | 0.6 | 2.7 | 1.3 | XP_004525703 | PREDICTED: protein extra-macrochaetae-like [Ceratitis capitata] |
| 10944 | 2255 | 8.96 | 47.5 | 43.7 | 24.7 | 38.6 | 1.6 | 5.2 | 6.1 | 4.3 | XP_004522803 | PREDICTED: ATP-dependent RNA helicase DHX8-like isoform X2 [Ceratitis capitata] |
| 11509 | 4304 | 15.09 | 25.9 | 30.9 | 14.8 | 23.8 | 0.6 | 2.1 | 2.0 | 1.6 | XP_004519049 | PREDICTED: la-related protein CG11505-like isoform X5 [Ceratitis capitata] |
| 11846 | 281 | 11.63 | 31.6 | 23.9 | 15.3 | 23.6 | 0.8 | 0.0 | 5.3 | 2.0 | EFN64458 | PiggyBac transposable element-derived protein 4 [Camponotus floridanus] |
| 11944 | 1366 | 3.56 | 36.8 | 29.8 | 51.3 | 39.3 | 6.6 | 13.6 | 12.9 | 11.1 | XP_004533749 | PREDICTED: ATPase ASNA1 homolog [Ceratitis capitata] |
| 12135 | 1736 | 3.95 | 63.7 | 61.8 | 89.9 | 71.8 | 13.9 | 20.8 | 19.8 | 18.2 | XP_004521155 | PREDICTED: uncharacterized protein LOC101455781 [Ceratitis capitata] |
| 12338 | 378 | 91.93 | 35.1 | 43.7 | 75.1 | 51.3 | 0.9 | 0.3 | 0.5 | 0.6 | XP_004526058 | PREDICTED: UDP-glucuronosyltransferase 2B15-like [Ceratitis capitata] |
| 12582 | 669 | 33.81 | 39.1 | 37.0 | 21.9 | 32.7 | 0.2 | 0.2 | 2.6 | 1.0 | XP_004213206 | PREDICTED: piggyBac transposable element-derived protein 4-like [Hydra magnipapillata] |
| 13120 | 350 | 20.91 | 52.8 | 38.2 | 23.5 | 38.2 | 1.7 | 0.2 | 3.6 | 1.8 | XP_004522172 | PREDICTED: zygotic gap protein knirps-like isoform X1 [Ceratitis capitata] |
| 13344 | 1521 | 29.64 | 54.3 | 55.0 | 27.7 | 45.7 | 0.3 | 0.0 | 4.3 | 1.5 | XP_004527365 | PREDICTED: segmentation protein Runt-like isoform X1 [Ceratitis capitata] |

<table>
<tr><th></th><th></th><th></th><th colspan="8">Normalised RPKM values</th><th colspan="2">Top BLAST Hit</th></tr>
<tr><th></th><th></th><th></th><th>male</th><th colspan="2">female</th><th>“late”</th><th>male</th><th colspan="2">female</th><th>“early”</th><th></th><th></th></tr>
<tr><th>Contig ID</th><th>Contig length</th><th>Fold Change</th><th>BJ2</th><th>BJ3</th><th>BJ4</th><th>mean</th><th>BJ6</th><th>BJ7</th><th>BJ8</th><th>mean</th><th>Accession (E-value)</th><th>Description</th></tr>
<tr><td>13492</td><td>551</td><td>5.8</td><td>64.6</td><td>56.9</td><td>38.3</td><td>53.3</td><td>5.4</td><td>11.7</td><td>10.5</td><td>9.2</td><td>XP_004530499</td><td>PREDICTED: probable serine/threonine-protein kinase yakA-like [Ceratitis capitata]</td></tr>
<tr><td>13588</td><td>1070</td><td>38.38</td><td>40.3</td><td>40.0</td><td>20.7</td><td>33.7</td><td>0.3</td><td>0.5</td><td>1.8</td><td>0.9</td><td>NP_001266341</td><td>>gi|214028056|gb|ABW97511.1| alpha-esterase 7 [Ceratitis capitata]</td></tr>
<tr><td>13723</td><td>701</td><td>32.3</td><td>46.3</td><td>41.8</td><td>21.9</td><td>36.7</td><td>0.4</td><td>0.8</td><td>2.2</td><td>1.1</td><td>NP_001266341</td><td>>gi|214028056|gb|ABW97511.1| alpha-esterase 7 [Ceratitis capitata]</td></tr>
<tr><td>14900</td><td>337</td><td>2.72</td><td>57.2</td><td>50.2</td><td>34.5</td><td>47.3</td><td>13.1</td><td>19.3</td><td>19.8</td><td>17.4</td><td>XP_004522151</td><td>PREDICTED: DNA-directed RNA polymerase III subunit RPC4-like isoform X2 [Ceratitis capitata]</td></tr>
<tr><td>15106</td><td>2560</td><td>2.59</td><td>46.9</td><td>41.5</td><td>32.6</td><td>40.3</td><td>10.9</td><td>17.5</td><td>18.4</td><td>15.6</td><td>XP_004531488</td><td>PREDICTED: protein sly1 homolog [Ceratitis capitata]</td></tr>
<tr><td>15168</td><td>2974</td><td>135.45</td><td>25.9</td><td>18.0</td><td>40.4</td><td>28.1</td><td>0.2</td><td>0.1</td><td>0.3</td><td>0.2</td><td>AAN87269</td><td>ORF [Drosophila melanogaster]</td></tr>
<tr><td>15265</td><td>1176</td><td>49.3</td><td>28.3</td><td>27.8</td><td>16.6</td><td>24.3</td><td>0.2</td><td>0.1</td><td>1.2</td><td>0.5</td><td>NP_476730</td><td>>gi|7295755|gb|AAF51058.1| sloppy paired 1 [Drosophila melanogaster]</td></tr>
<tr><td>15562</td><td>1261</td><td>5.89</td><td>37.0</td><td>36.2</td><td>54.4</td><td>42.6</td><td>2.8</td><td>10.0</td><td>8.8</td><td>7.2</td><td>XP_004534447</td><td>PREDICTED: uncharacterized protein LOC101461354 isoform X1 [Ceratitis capitata]</td></tr>
<tr><td>15731</td><td>492</td><td>12.97</td><td>23.8</td><td>19.4</td><td>13.1</td><td>18.8</td><td>0.6</td><td>0.0</td><td>3.7</td><td>1.5</td><td>XP_002060047</td><td>GJ15515 [Drosophila virilis] >gi|194141845|gb|EDW58258.1| GJ15515 [Drosophila virilis]</td></tr>
<tr><td>16200</td><td>360</td><td>4.2</td><td>31.5</td><td>23.3</td><td>22.1</td><td>25.6</td><td>1.4</td><td>8.4</td><td>8.6</td><td>6.1</td><td>XP_004520506</td><td>PREDICTED: cyclin-H-like [Ceratitis capitata]</td></tr>
<tr><td>16518</td><td>570</td><td>4.11</td><td>47.9</td><td>42.0</td><td>69.0</td><td>52.9</td><td>12.3</td><td>10.1</td><td>16.2</td><td>12.9</td><td>XP_004530540</td><td>PREDICTED: host cell factor-like [Ceratitis capitata]</td></tr>
<tr><td>16546</td><td>2732</td><td>6.16</td><td>36.9</td><td>31.8</td><td>21.1</td><td>29.9</td><td>2.8</td><td>5.0</td><td>6.8</td><td>4.9</td><td>XP_004537763</td><td>PREDICTED: inhibitor of growth protein 3-like [Ceratitis capitata]</td></tr>
<tr><td>16608</td><td>2306</td><td>6.82</td><td>36.6</td><td>31.5</td><td>53.4</td><td>40.5</td><td>2.3</td><td>8.8</td><td>6.7</td><td>5.9</td><td>XP_004525046</td><td>PREDICTED: protein DDI1 homolog 2-like isoform X1 [Ceratitis capitata]</td></tr>
<tr><td>17055</td><td>222</td><td>14.79</td><td>142.4</td><td>107.2</td><td>71.8</td><td>107.1</td><td>4.4</td><td>0.0</td><td>17.3</td><td>7.2</td><td>XP_004523465</td><td>PREDICTED: protein spaetzle-like isoform X2 [Ceratitis capitata]</td></tr>
<tr><td>18109</td><td>924</td><td>3.62</td><td>26.3</td><td>21.0</td><td>20.2</td><td>22.5</td><td>2.5</td><td>7.8</td><td>8.3</td><td>6.2</td><td>XP_004535038</td><td>PREDICTED: carbonic anhydrase 2-like isoform X1 [Ceratitis capitata]</td></tr>
<tr><td>19290</td><td>698</td><td>5.6</td><td>51.7</td><td>43.4</td><td>68.8</td><td>54.6</td><td>1.6</td><td>14.2</td><td>13.5</td><td>9.8</td><td>XP_004518396</td><td>PREDICTED: DNA polymerase V-like [Ceratitis capitata]</td></tr>
<tr><td>20291</td><td>270</td><td>9.57</td><td>36.1</td><td>37.7</td><td>19.7</td><td>31.2</td><td>2.1</td><td>3.4</td><td>4.2</td><td>3.3</td><td>XP_004522802</td><td>PREDICTED: ATP-dependent RNA helicase DHX8-like isoform X1 [Ceratitis capitata]</td></tr>
<tr><td>20341</td><td>1943</td><td>7.31</td><td>23.3</td><td>15.8</td><td>27.4</td><td>22.2</td><td>1.1</td><td>4.1</td><td>3.9</td><td>3.0</td><td>XP_004526995</td><td>PREDICTED: serine-rich adhesin for platelets-like isoform X1 [Ceratitis capitata]</td></tr>
<tr><td>20639</td><td>1028</td><td>17</td><td>38.7</td><td>36.7</td><td>63.1</td><td>46.2</td><td>2.4</td><td>2.3</td><td>3.5</td><td>2.7</td><td>XP_004536159</td><td>PREDICTED: kelch-like protein 5-like [Ceratitis capitata]</td></tr>
<tr><td>23008</td><td>1377</td><td>4.88</td><td>35.4</td><td>28.2</td><td>47.8</td><td>37.1</td><td>2.0</td><td>10.8</td><td>10.1</td><td>7.6</td><td>XP_004518396</td><td>PREDICTED: DNA polymerase V-like [Ceratitis capitata]</td></tr>
</table>

**Table E.5.** A list of the *Bactrocera jarvisi* contigs up-regulated in older embryos with BLAST matches to *C. capitata* and *Drosophila* genes (62 contigs). These were cross-referenced to FlyBase. The molecular and biological function of these genes and the expression profile in *D. melanogaster* are listed. Many genes have transcription or translation regulatory functions, but are transcribed in both males and females (see Table E.4).

| | Closest *Drosophila* match | | | |
|---|---|---|---|---|
| **contig ID** | **Annotation ID** | **Gene name** | **Molecular or biological function** | **Expression profile** |
| 36 | CG7163 | mkg-p, monkey-king protein | nucleotidyltransferase activity | maternal |
| 930 | CG10161 | eiF-3p66, Eukaryotic initiation factor 3 p66 subunit | translation initiation factor activity | |
| 1038 | CG9684 | | mRNA splicing, via spliceosome | 0-6h, adult female |
| 1228 | CG7483 | eIF4AIII | ATP binding; protein binding; translation initiation factor activity; ATP-dependent RNA helicase activity | 0-18h |
| 1817 | CG40218 | Yeti | kinesin binding | 0-6h, 12-18h |
| 2314 | CG7878 | | ATP binding; RNA binding; ATP-dependent RNA helicase activity; helicase activity | 0-6h, maternal, adult females. |
| 2876 | CG2848 | Trn-SR, Transportin-Serine/Arginine rich | Ran GTPase binding; protein binding | 0-12h, maternal, adult female reproductive system. |
| 3007 | CG12205 | Bsg25A, Blastoderm-specific gene 25A | sequence-specific DNA binding; chromatin insulator sequence binding | Detected during nuclear cycle 10-13 |
| 3009 | CG12205 | Bsg25A, Blastoderm-specific gene 25A | sequence-specific DNA binding; chromatin insulator sequence binding | Detected during nuclear cycle 10-13 |
| 3156 | CG11958 | Cnx99A, Calnexin 99A | calcium ion binding; unfolded protein binding | 0-18h |
| 3353 | CG3227 | insv, insensitive | transcription corepressor activity | 0-6h, maternal |
| 3466 | CG8384 | gro, groucho | RNA polymerase II transcription corepressor activity; transcription corepressor activity; transcription factor binding; protein binding; repressing transcription factor binding; HMG box domain binding | 0-6h, maternal |
| 3515 | CG6605 | Bicaudal D | protein binding | 0-6h and adult females, maternal, rapidly degraded stage 4-6 |
| 3878 | CG3048 | TNF receptor-associated factor 4 | zinc-finger protein binding | 0-12h, stage4-6 |
| 4660 | CG5358 | Arginine methyltransferase 4 (Art4) | histonearginine N-methyltranferase activity | 0-6h and adult female, maternal |
| 4679 | CG9854 | hrg, hiragi | polynucleotide adenylyltransferase activity; RNA binding | 0-6h, maternal, rapidly degraded at stage 4-6 |
| 4783 | CG1839 | Fbxl4, Fbox and leucine-rich-repeat gene4 | | 0-12h |
| 4784 | CG1839 | Fbxl4, Fbox and leucine-rich-repeat gene4 | | 0-12h |
| 4801 | CG18259 | | nucleotide binding; nucleic acid binding | 0-12h |
| 4934 | CG3227 | insv, insensitive | transcription corepressor activity | 0-6h, maternal |
| 4984 | CG3227 | insv, insensitive | transcription corepressor activity | 0-6h, maternal |
| 5000 | CG5490 | Toll | protein binding | embryo stage 14-16 and pupae |

| contig ID | Closest *Drosophila* match<br>Annotation ID | Gene name | Molecular or biological function | Expression profile |
|---|---|---|---|---|
| 6083 | CG12205 | Bsg25A, Blastoderm-specific gene 25A | sequence-specific DNA binding; chromatin insulator sequence binding | Detected during nuclear cycle 10-13 |
| 6152 | CG9775 | | | 0-12h, maternal |
| 6356 | CG9883 | Elba2 Early boundary activity 2 | chromatin insulator sequence binding; sequence specific binding | 0-6h, maternal, present adult ovary, |
| 6406 | CG13713 | | regulation of localization | 0-6h |
| 6564 | CG4184 | MED15, Mediator complex subunit 15 | regulation of transcription from RNA polymerase II promoter; transcription initiation from RNA polymerase II promoter | 0-6h, maternal |
| 7120 | CG14064 | beat-VI | heterophilic cell-cell adhesion | 12-24h |
| 7485 | CG8241 | pea, peanuts | ATP binding; RNA binding; regulation of alternative mRNA splicing, via spliceosome | 0-6h, adult ovary |
| 7654 | CG4878 | eIF3-S9 | mRNA binding; translation initiation factor activity; nucleotide binding; translation initiation factor binding | maternal |
| 7816 | CG1639 | l(1)10Bb, lethal (1) 10Bb | | 0-6h, maternal |
| 8666 | CG10360 | ref(2)P, refractory to sigma P | predicted zincion binding | 0-6h, maternal, male and female adults, |
| 8873 | CG6197 | | regulation of alternative mRNA splicing, via spliceosome | 0-6h, adult ovary |
| 8928 | CG7581 | Bub3 | mitotic cell cycle checkpoint,  mitotic spindle organisation and mitotic spindle assembly checkpoint | 0-12h, maternal |
| 10611 | CG1007 | emc, extra-macrochaetae | protein binding transcription factor activity, protein heterodimerisation activity, predict transcription corepressor activity | 0-18h, maternal |
| 10944 | CG8241 | pea, peanuts | ATP binding; RNA binding; regulation of alternative mRNA splicing, via spliceosome | 0-6h, adult ovary |
| 11509 | CG11505 | | 1 HTH La-type RNA-binding domain, predicted nucleic acid binding, nucleotide binding | 0-6h, adult ovary and testis |
| 11944 | CG1598 | | ATP binding; arsenite-transmembrane transporting ATPase activity | 0-6h, maternal |
| 12135 | CG9684 | | mRNA splicing, via spliceosome | 0-6h, adult female |
| 12338 | CG15661 | | glucuronosyltransferase activity | 6-24h |
| 13120 | CG4717 | kni, knirps | repressing transcription factor binding, sequence-specific DNA binding transcription factor activity | 0-6h, maternal |
| 13344 | CG1849 | runt | sequence-specific DNA binding transcription factor activity; ATP binding | 0-12h, zygotic |
| 13492 | CG12701 | zelda?vfl, vielfaltig (also zld) | sequence-specific DNA binding; metal ion binding; transcription regulatory region sequence-specific DNA binding | 0-12h, maternal, adult ovary |
| 13588 | CG1112 | a-Est7, a-Esterase-7 | carboxylesterase activity | embryo stage 11, adult male |

| contig ID | Closest *Drosophila* match<br>Annotation ID | Gene name | Molecular or biological function | Expression profile |
|---|---|---|---|---|
| 13723 | CG1112 | a-Est7, a-Esterase-7 | carboxylesterase activity | embryo stage 11, adult male |
| 14900 | CG5147 | | DNA binding; DNA-directed RNA polymerase activity | 0-18h |
| 15106 | CG3539 | Slh, SLY-1 homologous | predicted SNARE binding | maternal |
| 15168 | na | roo/ORF, ORF | predicted RNA-binding, RNA-directed DNA polymerase activity, zinc ion binding | |
| 15265 | CG16738 | slp1, sloppy paired 1 | sequence-specific DNA binding transcription factor activity | 0-12h, first expressed stage 4 |
| 15562 | CG40191 | | protein kinase binding | 6-12h |
| 15731 | CG13713 | | | 0-6h, zygotic |
| 16200 | CG7405 | cycH, Cyclin H | RNA polymerase II carboxy-terminal domain kinase activity; protein kinase binding | 0-12h, adult female |
| 16518 | CG1710 | Hcf, Host cell factor | sequence-specific DNA binding transcription factor activity; histone acetyltransferase activity; chromatin binding; transcription coactivator activity | embryogenesis, maternal, adult female |
| 16546 | CG42638 | | Polycystin cation channel | embryogenesis, adults |
| 16608 | CG4420 | rngo. Rings lost | proteasome binding; ubiquitin binding; aspartic-type endopeptidase activity | 0-12h, maternal, adult female |
| 17055 | CG6134 | spz, spatzle | morphogen activity; protein homodimerization activity; cytokine activity; growth factor activity; Toll binding | 0-6h, maternal, female adult |
| 18109 | CG11284 | | carbonate dehydratase activity | 0-6h, adult females |
| 19290 | CG6189 | l(1)1Bi, lethal (1) 1Bi | DNA binding; DNA-directed DNA polymerase activity | embryogenesis, early larval, adult female |
| 20291 | CG8241 | pea, peanuts | ATP binding; RNA binding; regulation of alternative mRNA splicing, via spliceosome | 0-6h, adult ovary |
| 20341 | CG33232 | | actin binding | 0-6h, adult female reproductive system |
| 20639 | CG7210 | kel, kelch | actin binding | 0-6, 18-24h, maternal |
| 23008 | CG6189 | l(1)1Bi, lethal (1) 1Bi | DNA binding; DNA-directed DNA polymerase activity | embryogenesis, early larval, adult female |